GENETIC AND EPIGENETIC MECHANISMS OF TUMORIGENESIS IN LYNCH AND LYNCH-LIKE SYNDROME

by

Noora Porkka

Department of Medical and Clinical Genetics
Medicum
Faculty of Medicine
University of Helsinki
Finland

Doctoral School of Health Sciences
Doctoral program in Integrative Life Science

ACADEMIC DISSERTATION

Doctoral thesis, to be presented for public examination with permission of
the Faculty of Medicine of the University of Helsinki
in the Lecture Hall 2, Biomedicum I, Helsinki
on the 17th of April 2020, at 12 pm

Supervised by

Professor Päivi Peltomäki, M.D., Ph.D.
Department of Medical and Clinical Genetics
Medicum, University of Helsinki, Finland

Docent Taina Nieminen, Ph.D.
Department of Medical and Clinical Genetics
Medicum, University of Helsinki, Finland

Reviewed by

Docent Katri Pylkäs, Ph.D.
Cancer and Translational Medicine Research Unit
University of Oulu, Finland

Docent Synnöve Staff, Ph.D.
Department of Obstetrics and Gynaecology
University of Tampere, Finland

Opponent

Professor Arto Mannermaa
Institute of Clinical Medicine
University of Eastern Finland, Finland

Custos

Professor Päivi Peltomäki

ISBN xx-xx-xx-xxxx-x (PRINT)
ISBN xxx-xxx-xx-xxxx-x (PDF)
ISSN xxxx-xxxx
PAINO Oy - City, Finland 2020

XX

ABSTRACT

Lynch syndrome is the most prevalent cancer predisposition syndrome that causes significantly increased lifetime risk of cancer in multiple organs such as colorectum, endometrium and ovarium. The predisposition is caused by germline mutations in DNA mismatch repair (MRR) genes *MLH1*, *MSH2*, *MSH6*, and *PMS2*. Lynch-like syndrome colorectal tumors, like Lynch syndrome tumors, are MMR-deficient. Nevertheless, Lynch-like syndrome tumors do not bear germline mutations in MMR genes nor methylation of the promoter regions of MMR genes that would explain the deficiency. Instead, MMR deficiency in majority of Lynch-like syndrome tumors is caused by two somatic mutations in the MMR genes. Dysfunctional MMR protein complex enables accumulation of mutations in the genome (mutator phenotype) and, eventually, microsatellite instability and cancer.

The reasons behind organ selectivity in MMR-deficient tumors are unknown, and whether breast cancer is part of the Lynch syndrome spectrum is under debate. The germline mutations in MMR genes are well studied, but molecular characteristics of Lynch syndrome tumors remain to be studied further. Lynch-like syndrome tumors remain less well characterized: besides the double somatic MMR mutations as a cause of MMR deficiency, their molecular and clinicopathological features as well as their incidence in the population remain poorly known. Currently, the only possibility to diagnose Lynch-like syndrome is by ruling out the possibility of germline mutations in MMR genes (Lynch syndrome) or methylation of the promoter regions of MMR genes. Synchronous ovarian and endometrial carcinomas are common in Lynch syndrome; whether they are of same origin (metastatic cancer) or two independently developed primary cancers remains to be resolved. We aimed to characterize the epigenetic and somatic mutation profiles in Lynch syndrome representing colorectal, ovarian, endometrial and breast carcinomas, and to identify new unique features that could be used in evaluating cancer risk, diagnosis and targeted treatment.

We used targeted high-throughput sequencing of 578 known cancer genes to investigate the somatic mutation profiles, and methylation-specific multiplex ligation-dependent

probe amplification to study the epigenetic profiles of the tumors. Non-synonymous somatic mutations were detected from sequencing data of the paired tumor and normal tissues to determine mutation signatures and identify potential driver genes. The data was also compared statistically between tumors of different origin, epigenetic status, and between breast carcinomas from Lynch syndrome mutation carriers and their known non-carrier family members.

We observed that Lynch- and Lynch-like syndrome tumors have unique somatic mutation and methylation profiles. We were able to link the methylator phenotype to high somatic mutation rates, and Lynch-like colorectal tumors to hypermethylated CpG island methylator phenotype (CIMP), which are novel findings. Our discovery of high mutation burden in genes associated with epigenetic regulation provides a new link between genetic and epigenetic factors in tumorigenesis. Genetic and epigenetic characterization of synchronous ovarian and endometrial carcinomas indicated shared origin, in analogy to sporadic cases. Molecular characteristics and especially mutational signatures of breast tumors of Lynch syndrome mutation carriers indicated that breast carcinoma is likely to be part of the Lynch syndrome tumor spectrum.

These findings bear potential clinical relevance since the molecular tumor profiles may be used in diagnosis and may guide tailored management of the patients. Many of the mutated genes are part of signaling routes to which targeted molecules either exist or can be developed.

Keywords: Lynch syndrome, Lynch-like syndrome, panel sequencing, epigenetics, somatic mutation profile, promoter methylation, colorectal cancer, ovarian cancer, breast cancer

TIIVISTELMÄ

Lynchin syndrooma on yleisin syövälle altistava perinnöllinen syndrooma, joka lisää muun muassa suolisto-, kohtu- ja munasarjasyövän riskiä. Alttiuden aiheuttavat perinnölliset mutaatiot DNA:n emäspariutumisvirheitä korjaavissa (engl. DNA mismatch repair, MMR) geeneissä *MLH1*, *MSH2*, *MSH6*, and *PMS2*. MMR-proteiineista koostuvan koneiston puutteellinen toiminta johtaa mutaatioiden kertymiseen genomiin ja lopulta mikrosatelliitti-epätasapainoon ja syövän syntyyn. Lynchin kaltaisen (engl. Lynch-like) syndrooman suolistokasvainten MMR-proteiinien puutos ei selity ituratamutaatioilla tai MMR-geenien promoottorialueiden hypermetylaatiolla. Tiedetään, että valtaosassa tapauksista Lynch-like -kasvainten MMR-geenien normaali toiminta on hiljentynyt kahden somaattisen eli hankitun mutaation seurauksena.

Kudosspesifisyyden syitä syövissä, joissa MMR-koneiston toiminta on puutteellista, ei vielä tunneta. Rintasyövän kuulumisesta Lynchin syndrooman syöpäspektriin on ristiriitaisia tutkimustuloksia, ja synkronisten munasarja- ja kohdunrungonsyövän alkuperästä (metastasoitunut kasvain vai itsenäisesti kehittyneet primaarikasvaimet) ei löydy julkaistua tutkimustietoa. MMR-geenien perinnölliset mutaatiot tunnetaan jo melko hyvin, mutta Lynchin syndrooman kasvainten muista molekulaarisista ominaisuuksista tarvitaan vielä lisää tutkimustietoa. Lynch-like -kasvaimista on niukemmin tutkimustietoa: kahden somaattisen MMR-mutaation lisäksi niiden molekulaarisista ja kliinispatologisista ominaisuuksista sekä kyseisen tautimuodon esiintyvyydestä populaatiotasolla tiedetään hyvin vähän. Toistaiseksi Lynch-like -syndrooma on mahdollista diagnosoida vain poissulkemalla Lynchin syndrooma ja MMR-geenien hypermetylaatio. Tavoitteenamme oli kartoittaa Lynchin ja Lynch-like -syndroomaan kuuluvien suolisto-, munasarja-, kohdunrungon ja rintasyöpien somaattisia ja epigeneettisiä muutoksia. Näiden piirteiden tunnistaminen eri syöpätyypeissä mahdollistaisi syövän riskin arvioinnin, tarkemman diagnosoinnin sekä kohdennetun hoidon kehittämisen.

Tutkimme somaattisia mutaatioita sekvensoimalla kasvainten DNA:ta paneelilla, joka kattaa 578 tunnettua syöpägeeniä, sekä niiden metylaatioprofiilia ns. MS-MLPA (eng.

methylation-specific multiplex ligation-dependent probe amplification) -tekniikan avulla. Tunnistimme ei-synonyymiset mutaatiot parittaisesta kasvain- ja normaalikudosten sekvenssidatasta määrittääksemme mutaatioprofiilit sekä syövänkehityksen ajurigeenit (engl. driver genes). Tilastolliset vertailut tehtiin eri kudosten ja metylaatiostatusten kesken. Lisäksi Lynchin syndrooman mutaationkantajilta peräisin olevia rintasyöpiä verrattiin rintasyöpiin sukulaisilta, jotka eivät olleet perineet sukunsa alttiusmutaatiota.

Havaitsimme, että Lynchin ja Lynch-like syndroomien syövillä on niille luonteenomaiset epigeneettiset ja somaattiset mutaatioprofiilit. Uusina löydöksinä osoitimme yhteyden yleistyneen hypermetylaatio- ja mutaatiotaipumuksen välillä, sekä yhteyden Lynch-like -kasvainten ja hypermetyloituneen fenotyypin välillä. Epigeneettiseen säätelyyn osallistuvien geenien sekä DNA:n korjausgeenien lisääntynyt mutaatioherkkyys liittää geneettiset ja epigeneettiset tekijät uudella tavalla toisiinsa syövänkehityksessä. Se, että synkroniset munasarjan ja kohdunrungon kasvaimet olivat geneettisesti ja epigeneettisesti yhteneväisiä, osoittaa, että kasvaimet ovat samaa alkuperää, eli edustavat metastasoitunutta kasvainta, kuten vastaavissa sporadisissa syövissä on aiemmin osoitettu. Lynchin syndrooman kantajien ja ei-kantajien rintasyöpien molekulaariset tutkimukset ja erityisesti mutaatioprofiilit toivat viitteitä siitä, että rintasyöpä kuuluu Lynchin syndrooman kasvainspektriin.

Tutkimus tuo tärkeää uutta tietoa Lynchin ja Lynch-like syndrooman syöpien syntymekanismeista sekä ominaispiirteistä, joita on mahdollista hyödyntää syövän diagnosoinnissa sekä kohdennettua hoitoa suunniteltaessa. Monet mutatoituneista geeneistä kuuluvat säätelyreitteihin, joihin on joko olemassa tai kehitettävissä kohdennettuja lääkemolekyylejä.

Avainsanat: Lynchin syndrooma, Lynch-like syndrooma, paneelisekvensointi, promoottorin metylaatio, somaattinen mutaatioprofiili, epigenetiikka, suolistosyöpä, munasarjasyöpä, rintasyöpä

TABLE OF CONTENTS

ABBREVIATIONS

MMR	DNA mismatch repair
LS	Lynch syndrome
MSI	microsatellite instability
LLS	Lynch-like syndrome
CIN	chromosomal instability
miRNA	micro-ribonucleic acid
CMMRD	constitutional MMR deficiency
AC	Amsterdam criteria
BG	Bethesda guidelines
LSRFi	The National Lynch syndrome registry of Finland
FAP	Familial adenomatous polyposis
IHC	immunohistochemistry
CRC	colorectal carcinoma
AD	adenoma
OC	ovarian carcinoma
EC	endometrial carcinoma
BC	breast carcinoma
NC-BC	non-carrier breast carcinoma
FFPE	formalin-fixed paraffin embedded
TMA	tissue microarray
MS-MLPA	Methylation-specific multiplex ligation-dependent probe amplification
MethyQESD	Methylation-quantification of endonuclease-resistant DNA
CCP	Comprehensive Cancer Panel
PCR	polymerase chain reaction
CIMP	CpG island methylator phenotype
FIMM	Institute for Molecular Medicine Finland
LOH	Loss of heterozygosity
Alt	variant allele reads
Ref	reference allele reads
T	Tumor
N	Normal
R	LOH ratio
SNV	single nucleotides variants
dMMR	MMR-deficient
pMMR	MMR-proficient
PV	pathogenic variant
DS	double somatic
HBOC	Hereditary Breast and Ovarian Cancer Syndrome
NGS	Next Generation Sequencing
VUS	variant of unknown significance
CNV	copy-number variation

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following original publications, which are referred to in the text by the Roman numerals I-IV:

I Porkka N, Valo S, Nieminen TT, Olkinuora A, Mäki-Nevala S, Eldfors S, Peltomäki P. Sequencing of Lynch syndrome tumors reveals the importance of epigenetic alterations. *Oncotarget*. 2017 Nov 14;8(64):108020-108030.

II Porkka N, Lahtinen L, Ahtiainen M, Böhm J, Kuopio T, Eldfors S, Mecklin J-P, Seppälä T, Peltomäki P. Epidemiological, clinical and molecular characterization of Lynch-like syndrome – A population-based study. *International Journal of Cancer*. 2019 Jul 1;145(1):87-98.

III Porkka N, Olkinuora A, Kuopio T, Ahtiainen M, Eldfors S, Almusa H, Mecklin J-P, Peltomäki P. Does breast carcinoma belong to the Lynch syndrome tumor spectrum? – Somatic mutational profiles vs. ovarian and colorectal carcinomas. *Manuscript.*

IV Niskakoski A, Pasanen A, Porkka N, Eldfors S, Lassus H, Renkonen-Sinisalo L, Kaur S, Mecklin J-P, Bützow R, Peltomäki P. Converging endometrial and ovarian tumorigenesis in Lynch syndrome: Shared origin of synchronous carcinomas. *Gynecol Oncol*. 2018 Jul;150(1):92-98.

In addition, some unpublished data are presented in this thesis.

INTRODUCTION

Cancer is one of the leading causes of morbidity and the second leading cause of death globally. Regardless of the extensive effort and money spent on research, cancer was reported to be the cause of death in 9.6 million cases in 2018. However, over 50% of the patients diagnosed with cancer in Finland are currently cured (Cancer Society of Finland, All about Cancer). Environmental factors such as UV-radiation, tobacco smoke and Western style diet are known risk factors for many cancers. Evolving diagnostic methods and the fact that people tend to live longer may also increase the incidence of diagnosed cancers.

Cancer is a group of distinct diseases caused by genetic and epigenetic changes, affecting nearly all the organs and cell types in the human body. Genetics is defined as information coded in the DNA sequence, whereas epigenetics refers to inherited patterns of gene expression (Esteller and Herman, 2002). These changes disrupt normal function of the cell allowing uncontrolled cell proliferation and metastasis into distant tissues and organs (Stratton et al., 2009).

Every cell in a human body carries a single genome coded into two sets of chromosomes present in the fertilized egg. In addition to the environmental carcinogens, normal cells acquire somatic mutations in a steady rate caused internal mutagenic agents (eg. radical oxygen species caused by normal metabolic functions) and by errors occurring during DNA replication (Stratton et al., 2009). DNA repair machinery of a normal cell repairs most of these damages. In some cancers, the acquisition rate of somatic mutations may be increased by for example deficiencies in DNA mismatch repair (MMR) machinery which causes accumulation of somatic mutations in the genome that may eventually lead to cancer (Jeggo et al., 2016). These cancer-promoting mutations can also be germline mutations which are inherited from the parents.

Lynch syndrome (LS) is the most common cancer predisposition syndrome in the world (Win et al., 2017). It is an autosomal dominant disorder caused by germline mutations in one of the DNA mismatch repair (MRR) genes *MLH1*, *MSH2*, *MSH6*, and *PMS2* (Peltomäki, 2005). Dysfunction of MMR machinery leads to microsatellite instability (MSI) that is characteristic to majority of LS tumors, whereas less common in sporadic tumors (Lotsari et al., 2012, Peltomäki, 2003). MMR genes comply with the Knudson's two-hit mechanism: the first hit causes cancer predisposition and the second commences tumorigenesis (Knudson, 1971).

The revised Bethesda guidelines determines colorectal, endometrial, stomach, ovarian, ureter, renal pelvis, brain, small bowel, and hepatobiliary tract cancers as part of the LS tumor spectrum, due to the increased relative risk compared to general population, and the generally lower age of onset (Umar et al., 2004). Debate is ongoing whether breast cancer is part of LS tumor spectrum. Breast cancer is the most common cancer in women worldwide, but guidelines recommending increased breast cancer screening for women with LS are currently lacking.

Besides the germline mutations in MMR genes, little is known about the somatic mutational profiles or epigenetic profiles of the LS tumors, or the reasons for organ selectivity in germline mutation carriers. Only recently, Lynch-like syndrome (LLS) was established as a consistent group of MMR-deficient tumors by identification of double somatic mutations as the mechanism associated with majority of the cases (approximately up to 70%) (Pearlman et al., 2019). The mechanisms explaining the remaining cases, and the somatic and epigenetic profiles, are nevertheless, unknown.

Because Lynch syndrome is a well-established cancer syndrome, it provides a good model for both LS tumors and carcinogenesis in general. It serves as a valuable model to investigate the multistep tumorigenesis of colon and other organs: every cell of a LS mutation carrier bears the strong predisposing defect, but the tumor development requires an additional hit to occur in a target tissue. Discovering new unique somatic

features of LS and LLS tumors could enhance diagnostics and targeted treatment of these tumors.

REVIEW OF THE LITERATURE

1. Cancer in Finland

In 2017, almost 35 000 people were diagnosed with cancer in Finland, and the numbers are rising as they have been for over 60 years (tracked time) (Finnish Cancer registry). Two thirds of the diagnosed patients are cured (Cancer Society of Finland, All about Cancer). In Finland, the most common cancers in women are breast, colorectal and lung cancer, and prostate, colorectal and lung cancer in men (**Table 1**). Almost 2.1 million people worldwide develop breast cancer every year, being the second most common cancer in the world, and in Finland nearly 5000 new cases are diagnosed annually (Finnish Cancer registry, International Agency for Research on Cancer). Prostate cancer is the world's fourth most common cancer with nearly 1.3 million new diagnoses every year (International Agency for Research on Cancer), whereas it is the most common cancer in males in Finland (**Table 1**) (International Agency for Research on Cancer). Colorectal cancer is common in Western populations (Dermadi et al., 2014). It is the second most common cancer in both males and females in Finland with over 3000 new cases diagnosed annually (Finnish Cancer registry). Lung cancer is the third most common cancer in Finland in both women and men, but it is the most common cancer in the world with nearly 2.1 million new cases annually (**Table 1** (International Agency for Research on Cancer)). Ovarian cancer is less common in Finland, with only 442 patients diagnosed in 2017, and the incidence during the last ~30 years has remained somewhat even, unlike in many other cancers (Finnish Cancer registry).

Table 1. Statistics of cancer in Finland

Female					
#	Tissue	New cases	Deaths	Incidence*	Relative survival rate**
1.	Breast	4947	923	167.18	91 %
2.	Colorectal	1575	650	49.29	67 %
3.	Lung	999	780	31.61	19 %
...					
	Ovary	442	315	14.51	42 %
Male					
#	Tissue	New cases	Deaths	Incidence*	Relative survival rate**
1.	Prostate	5444	912	212.05	92 %
2.	Colorectal	1784	718	70.49	64 %
3.	Lung	1712	1502	67.11	12 %

*Relative per 100 000, age adjusted (Finland 2014)
**Relative percentage of the patients living five years after diagnosis (Surveillance period 2015-2017). Age adjusted.
This data is based on the Finnish Cancer Registry (Finnish Cancer Registry)

2. Hallmarks of cancer

Cancer is a genetic and epigenetic disease encompassing a group of distinct diseases emerging from almost all organs and cell types in the human body. This group of diseases is characterized by uncontrolled cell proliferation and metastasis, allowing cells to break through the tissue boundaries and invade to distant tissues and organs (Esteller and Herman, 2002, Stratton et al., 2009). The evolution of a normal cell into a cancerous cell occurs by multiple changes in DNA sequence, chromosome rearrangements and aneuploidy, leading to genome instability. This enables cells to sustain proliferative signaling, evade growth suppression, resist controlled cell death (apoptosis), induce angiogenesis, activate invasion and metastasis, and acquire replicative immortality. These biological capacities are known as hallmarks of cancer (Hanahan and Weinberg, 2000). Two additional emerging hallmarks have been discovered: reprogramming of energy metabolism and evading immune destruction. In addition, the complexity and

impact of tumor microenvironment was acknowledged as contributor of hallmark traits and, hence, tumorigenesis (Hanahan and Weinberg, 2011).

3. Genetics of tumorigenesis

Cancer develops through Darwinian natural selection: mutations enable adaptation to the changes in the surrounding microenvironment. Deleterious mutations may be excluded, whereas certain mutations may offer the cell an advantage to proliferate and survive more efficiently. The latter, however, occasionally results into development of cancer when the advantageous mutations allow a single cell to escape controlled cell proliferation and eventually metastasize (Stratton et al., 2009).

Every cell in the human body carries a diploid genome of the fertilized egg. On top of the germline mutations that are inherited from the parents, normal cells acquire mutations on a steady rate by external and internal mutagenic agents as well as errors occurring during the intrinsic DNA replication. These acquired mutations, called somatic mutations, differ cells from the progenitor fertilized egg and their parental cells, and are passed on to the daughter cells (Stratton et al., 2009). Tobacco smoke and ultraviolet radiation are examples of external carcinogens, whereas radical oxygen species from normal metabolic functions are classified as intrinsic mutation causing factors (Stratton et al., 2009). DNA repair machinery of a normal cell repairs most of these damages. The acquisition rate of somatic mutation may be increased in some cancers for example due to deficiencies in DNA mismatch repair (MMR) genes. This causes accumulation of single nucleotide changes and small insertions and deletions in the genome, causing a ‘mutator phenotype’ and, eventually, leading to MSI and cancer (Jeggo et al., 2016).

3.1 Driver and passenger mutations

Somatic mutations are classified as driver and passenger mutations according to their contribution to tumorigenesis. Driver mutations occur in cancer genes (Catalogue of somatic mutations in cancer (COSMIC)) that, according to their name, drive cancer development enabling growth advantage over the normal cells. Driver mutations are therefore positively selected during the course of cancer evolution (Stratton et al., 2009). Mutations not promoting cancer development are termed passenger mutations. They can occur relatively randomly around the genome and their effect on normal gene function may be neutral. Passenger mutations likely exist in an ancestor of a cancer cell solely by chance at the time the cell acquires a driver mutation turning it into a cancerous cell, without contributing to tumorigenesis (Vogelstein et al., 2013).

3.2 Oncogenes and tumor suppressor genes

Cancer-associated genes are typically classified as oncogenes and tumor suppressor genes. A normal form of an oncogene is called proto-oncogene which acts in important cellular functions such as cell growth, proliferation, cell survival and apoptosis (Croce, 2008). When aberrantly activated by for example a mutation, proto-oncogene is transferred into an active oncogene, which promotes growth and hence participates in driving tumorigenesis (Vogelstein et al., 2013). Tumor suppressor genes, contradictory to oncogenes, gain their tumor-promoting potential through inactivation of gene function (Vogelstein et al., 2013). In normal cells, they bear growth limiting potential and act as gatekeepers and caretakers by controlling cell division, DNA damage repair and apoptosis. Inactivation by mutation enables uncontrolled cell division, deficient DNA repair leading to accumulation of errors in the genome, and insufficient apoptosis (Sun and Yang, 2010). Mutations in oncogenes genes are typically dominant in their effect. In other words, one mutated allele of the gene is sufficient to promote carcinogenesis. (Stratton et al., 2009). Mutations in tumor suppressor genes, however, are usually

recessive in their effect on cellular level. This means that one normal allele is sufficient to retain normal control of cellular functions, and thus follows the Knudson two-hit hypothesis. According to Knudson two-hit hypothesis (Knudson, 1971) two mutations, genetic hits, are required to inactivate the gene in question. When one defective allele has been inherited in germline, only one somatic hit is required for inactivation and hence the carrier individuals tend to develop the disease at a younger age. Majority of cancer promoting genes are tumor suppressor genes, in which mutations cause defective proteins resulting in tumorigenesis (Rahman, 2014).

3.3 Genomic instability

Genomic instability is a common feature in human cancers that drives tumorigenesis by generating heterogeneity within the tumor cells inside the tumor (Andor et al., 2016). This intratumor heterogeneity provides genetic diversity against natural selection, enabling a tumor to efficiently adapt and survive. Chromosomal instability (CIN), a form of genomic instability, is a frequent feature of sporadic cancer but the mechanisms behind it remain to be studied further (Armaghany et al., 2012). CIN is a hallmark of tumors that evolve through the classical pathway, in which oncogenes such as *KRAS* and tumor suppressor genes such as *APC* or *TP53* are affected (Fearon and Vogelstein, 1990). Alterations associated with CIN are large-scale chromosomal rearrangements such as deletions, duplications, translocations, amplifications, insertions of transposable elements, or even aneuploidy (gain or loss of entire chromosomes) (Pino and Chung, 2010). MSI, another form of genomic instability, is characterized as variation in lengths of repeated microsatellites (small insertions and deletions) due to errors in DNA replication (Boland et al., 1998). Microsatellites are short nucleotide repeats scattered throughout the genome, often located in non-coding regions. The length of the repeats is unique in each individual (Jeffreys et al., 1992, Li et al., 2002). MSI results from deficient DNA repair leading to accumulation of mutations in the genome (mutator phenotype) which drives malignant transformation (Loeb, 1991). MSI is, hence, a

hallmark of cancers, in which DNA repair mutations are inherited (Lynch syndrome), but it is also detected in a proportion of different sporadic cancers including colorectal, gastric, endometrial, and ovarian (endometrioid) tumors. In sporadic cases MSI is associated with silenced expression of *MLH1* by promoter methylation (Imai and Yamamoto, 2008, Vasen et al., 2015).

3.4 *POLD1* and *POLE* mutations

Alternative mechanism for mutator phenotype are mutations in the genes *POLD1* and *POLE* encoding for DNA proofreading enzymes, resulting in increased number of base substitutions (Jansen et al., 2016, Palles et al., 2013). Hypermutability by *POLD1* and *POLE* mutations are observed in colorectal (Elsayed et al., 2015, Valle et al., 2014), endometrial (Church et al., 2015), and breast cancers (Voutsadakis, 2019).

4. Epigenetics of tumorigenesis

Epigenetic changes refer to inherited changes in gene expression without changes in the primary DNA sequence, that are thus potentially reversible (Feinberg and Tycko, 2004). In normal cells, epigenetic mechanisms are rigorous mechanisms for regulating homeostasis of gene expression in the cells (Esteller and Herman, 2002). Forms of epigenetic modifications consist of histone modifications, non-coding RNA expression (micro-ribonucleic acid, miRNA), modifications of chromatin remodeling system, and DNA methylation (Peltomäki, 2012). The latter is a common form of epigenetic change observed in cancer. DNA methylation mainly appears as modification of cytosine residues in CpG dinucleotides in CpG-rich regions called CpG islands mainly located in gene promoters (Sharma et al., 2010). Elevated DNA methylation (hypermethylation) in the promoter region can inhibit function of tumor suppressor genes, facilitating cancer development. However, decreased methylation (hypomethylation) in the promoter

regions of tumor promoting oncogenes may forward tumorigenesis (Baylin et al., 1986, Cross and Bird, 1995, Feinberg and Tycko, 2004) (**Figure 1**). For example, the methylation of the promoter region of *MLHI* is a common mechanism associated with the loss of *MLH1* gene expression, associated with a proportion of MSI tumors. Promising new therapies have been developed to target the heritable but reversible nature of epigenetic changes, which provide a platform for future treatments (Qi et al., 2016). CpG island DNA methylation is also an important regulatory element during embryonic development, for example by playing a role in X-chromosome inactivation and gene imprinting (Jones, 2012).

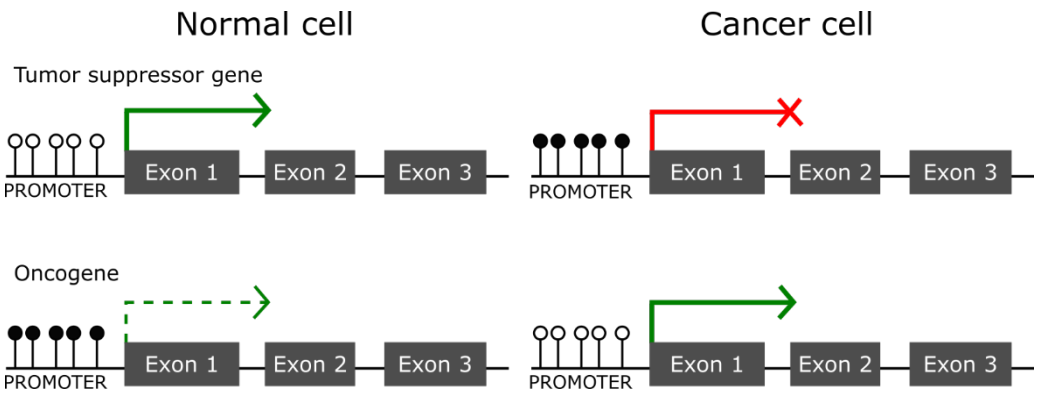


Figure 1. DNA methylation in cancer. The promoter hypermethylation (black circles, upper row) is a common mechanism for silencing tumor suppressor gene expression in cancer. Hypomethylation (white circles, lower row) of silenced oncogenes leads to gene activation. Green arrow indicates normally expressed gene, pointed green arrow indicates low gene expression, and red arrow with x indicates loss of gene expression. Modified from Peltomäki 2012 (Peltomäki, 2012).

5. Inherited cancer predisposition

Some individuals may have an inherited (germline) mutations predisposing to cancer development. Cancer itself is not contagious and it cannot be inherited from the parent, but the predisposition to cancer development can be inherited, which results in the lower age of onset. Heritability refers to the fraction of phenotypic variability attributable to genetic factors and ranges from 0% (all variability is due to environmental factors) to 100% (all variability is due to genetic factors). A large classical twin study conducted in the Nordic countries estimated that familial cumulative cancer risk was 32% and overall heritability 33% (Mucci et al., 2016). Significant heritability was observed in skin melanoma (58%), prostate cancer (57%), non-melanoma skin cancer (43%), ovarian cancer (39%), renal cancer (38%), breast cancer (31%), and cancer of corpus uteri (27%) (Mucci et al., 2016). In addition, heritability of 14% and 15% was observed in colon cancer and rectal and anal cancer, respectively (Mucci et al., 2016). A recent large genome-wide association study across six common cancers (breast, colorectal, head and neck, lung, ovarian and prostate cancers) indicated that common germline genetic mechanisms affect the development of solid tumors in different tissues (Jiang et al., 2019). Their single nucleotide polymorphism (SNP) heritability analysis estimated the following heritability values: approximately 12-17% for breast cancer, 7-12% for colorectal cancer, 5-15% for head and neck cancer, 6-10% for lung cancer, 2-5% for ovarian cancer, and 14-23% for prostate cancer (Jiang et al., 2019). These heritability values represent the familial aspect of cancer, indicating that more tissue-specific tumors arose within the families that would be expected by chance alone. However, heritability does not explain the reasons behind the phenomenon.

More than 110 genes have been discovered that are associated with hereditary cancer syndromes, of which the most common occur in genes involved in DNA repair (Rahman, 2014, Romero-Laorden and Castro, 2017). Reportedly, 2–8% of all colorectal cancers and 20% of colorectal cancers diagnosed before the age of 50 years are due to pathogenic germline variants in high risk cancer genes (AlDubayan et al., 2018, DeRycke et al.,

2013, Mork et al., 2015, Pearlman et al., 2017, Stoffel et al., 2018, Yurgelun et al., 2015). Classification of familial colorectal cancer syndromes are represented in **Figure 2**. Of ovarian cancers, as many as 18-26% represent a germline mutations in a cancer susceptibility gene (Hauke et al., 2019, Norquist et al., 2016, Toss et al., 2015), whereas in breast cancer the estimate is only 5-10% (Couch et al., 2014). These represent the hereditary form of cancer, since the underlying causative mechanism is known and is attributable to a single susceptibility gene with high-penetrance.

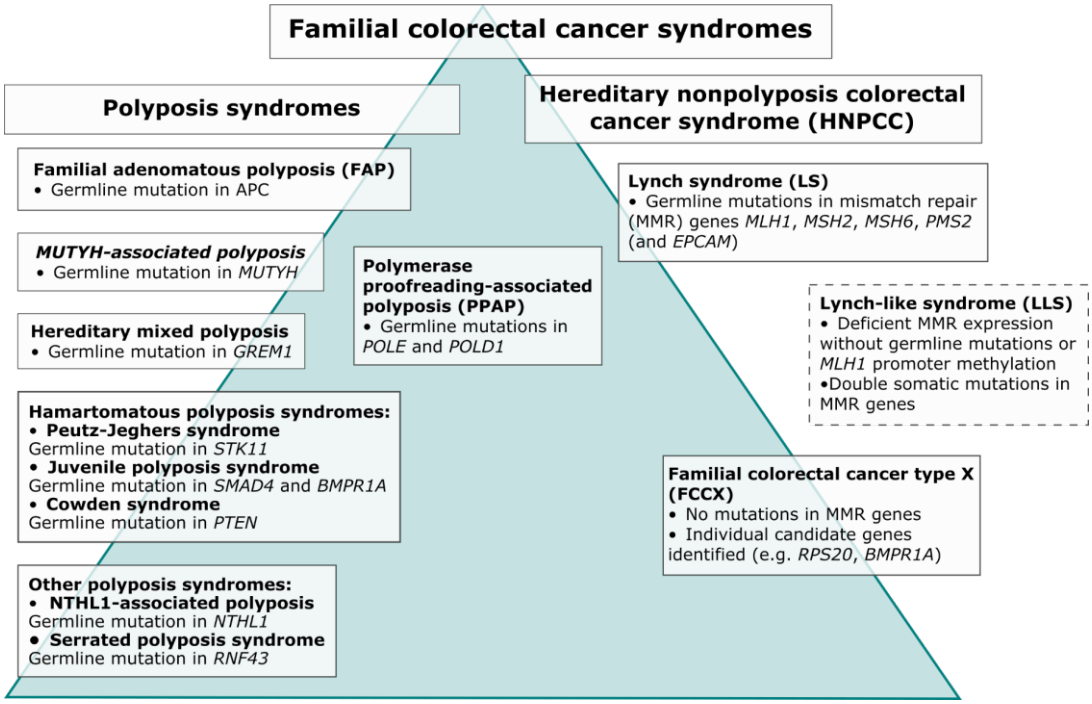


Figure 2. Familial colorectal cancer syndromes. Familial colorectal cancer syndromes are divided into polyposis syndromes (left) and non-polyposis syndromes (right). Lynch syndrome and Lynch-like syndrome are forms of non-polyposis syndromes, but Lynch-like syndrome is not hereditary and is hence represented in dashed line. Polymerase proofreading-associated polyposis (in the middle) is not categorized to either of the two categories. Based on information in Valle et al. (Valle et al., 2019).

5.1 Lynch syndrome

Lynch syndrome (LS) is the most common cancer predisposition syndrome and it is known to increase the risk of colorectal cancer and other epithelial malignancies, such as ovarian, endometrial, gastric, urinary tract and bladder, kidney, small bowel, and hepatobiliary tract tumors (Samadder et al., 2017, Watson and Riley, 2005). In addition, other cancers such as pancreatic and brain tumors occur more often in LS patients compared to average population (Lynch and de la Chapelle, 2003, Watson et al., 2008) (see **Figure 3**). Debate is ongoing on whether breast cancer is part of LS tumor spectrum (Aarnio et al., 1999, Buerki et al., 2012, Lotsari et al., 2012, Pande et al., 2012, Saita et al., 2018, Watson et al., 2008, Win et al., 2012). However, the reasons for organ selectivity in germline mutation carriers are unknown. Even though the incidence of an affected individual to develop cancer is high (75% in females and 58% in men before the age 70), the rate of survival in most cancers is higher compared to sporadic cancers (Moller et al., 2017b).

Up to 54% of female LS mutation carriers develop endometrial cancer and 24% ovarian cancer during their lifetime, making endometrial cancer as or even more prevalent cancer as CRC in female LS mutation carriers (Bonadona et al., 2011, Moller et al., 2017a). The histology of endometrial carcinomas in majority of LS cases is endometrioid (90%) (Rossi et al., 2017), whereas majority of epithelial ovarian carcinomas in LS are non-serous (77%) (Helder-Woolderink et al., 2016). Endometrial and ovarian carcinomas are diagnosed simultaneously in 10% of sporadic cases (Kelemen et al., 2017), in which recent studies indicate shared origin of the synchronous tumors (metastatic disease) (Anglesio et al., 2016, Schultheis et al., 2016). In LS, endometrial and ovarian carcinomas are diagnosed simultaneously in 20% of the cases (Rossi et al., 2017, Ryan et al., 2017). However, shared vs. independent origin of synchronous endometrial and ovarian carcinomas occurring in LS is unknown.

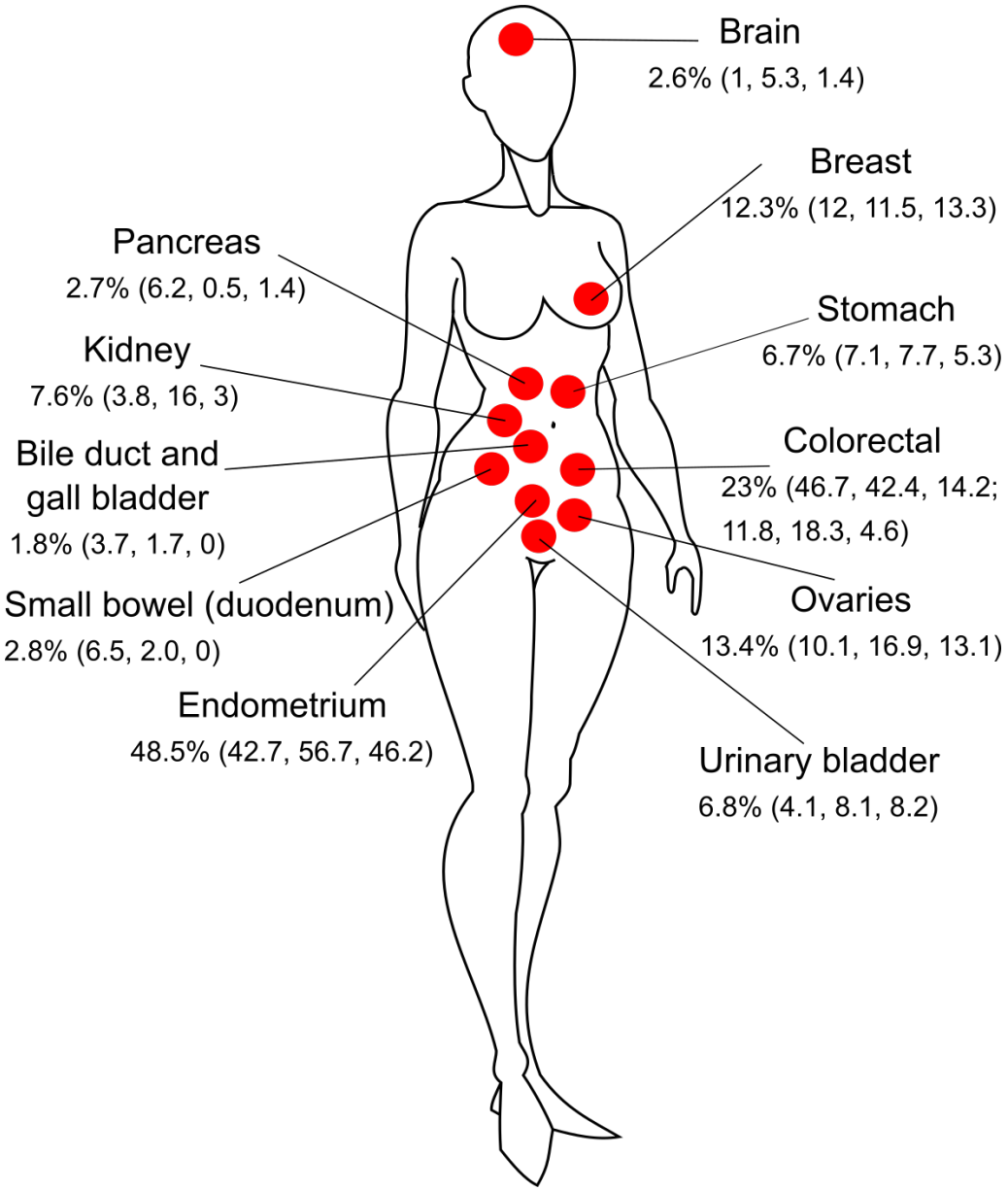


Figure 3. Cancer in Lynch syndrome. Average risk of different cancers in LS mutation carrier is calculated based on the risk reported per *MLH1*, *MSH2*, and *MSH6* (individual risks in parenthesis after average risk, respectively). Risk for breast, ovarian, and endometrial cancers are calculated in women only. The data is based on Møller *et al.* (Moller et al., 2018).

LS results from germline mutations in MMR genes *MLH1*, *MSH2*, *MSH6*, and *PMS2* (Thompson et al., 2014). Deletions in *EPCAM* associated with epigenetic silencing of *MSH2* are also linked to LS, but they are relatively rare. These truncating 3' end deletions cause epigenetic silencing of the neighbor *MSH2* gene in tissues in which EPCAM is expressed (Kuiper et al., 2011). An individual carrying a germline mutation in one of the MMR genes has a 50% risk to pass it to a child, representing the dominant characteristics of LS predisposing mutations. Extremely rare cases of constitutional MMR deficiency (CMMRD), severe form of childhood cancer syndrome, have been reported, in which a child inherits a defective MMR allele from both parents causing cancer already in early childhood (Wimmer et al., 2014). Numerous studies have been conducted regarding the germline mutations in MMR genes, but little is known about the somatic mutational profiles or epigenetic status of the LS tumors.

MMR proteins form a DNA repair complex that repairs mismatched bases and insertion/deletion loops in DNA that result from environmental factors or cellular processes, or that occur naturally during DNA replication (Liu et al., 2017) (**Figure 4**). Defects in MMR machinery cause accumulation of replication errors in the genome which can be detected as MSI, characteristic to most of the LS tumors (Peltomäki, 2001). When MMR complex is defective, small mono- and dinucleotide deletions and insertions accumulate in the genome leading to hypermutability and increased tumor heterogeneity advantageous to tumorigenesis (Jeggo et al., 2016). Inactivation can be caused by germline mutation in one of the four MMR genes combined with a second somatic hit inactivating the normal allele (Aaltonen et al., 1998). Inactivation or second hit can also be caused by hypermethylation of *MLH1* promoter region, which is the case in sporadic cancers (Jacob and Praz, 2002).

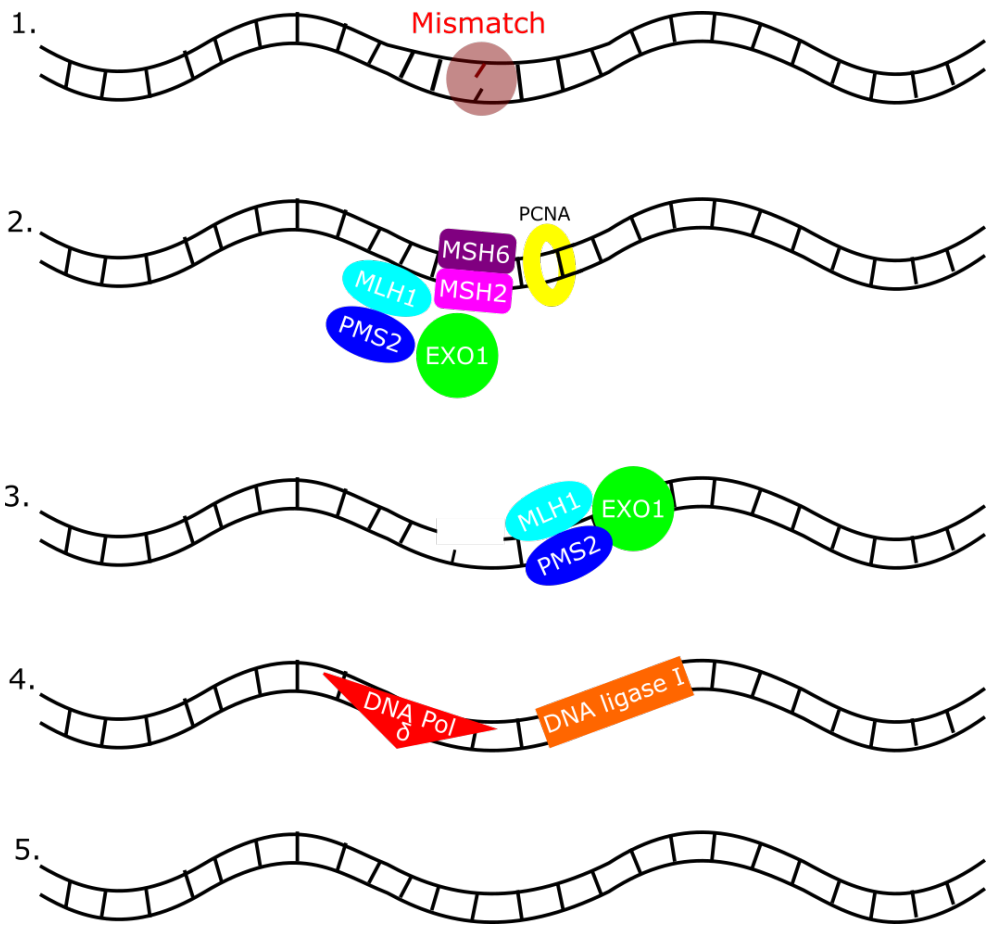


Figure 4. DNA mismatch repair. Simplified presentation of DNA mismatch repair pathway. 1. Mismatch recognition: Occurrence of single basepair mismatch in DNA synthesis. 2. Recognition of mismatch bases occurs by a complex of MSH6 and MSH2 heterodimers that combined form a hMutSα complex, together with Proliferating cell nuclear antigen (PCNA), both binding to mismatch site. MLH1 and PMS2 heterodimers combined form a hMutLα complex that is recruited to the site after recognition. 3. Together with exonuclease EXO1, hMutLα dissembles the mismatched DNA sequence. 4. DNA polymerase δ (DNA Pol δ) synthesizes a new complementary DNA strand, which is ligated with the undamaged strand by DNA ligase I. 5. Newly repaired DNA strand. Based on information in Liu et al. (Liu et al., 2017).

Before the genetic mechanisms of Lynch syndrome were established, Amsterdam criteria (AC) and Bethesda guidelines (BG), their variations AC II and revised BG,

together with the patient's pedigree and family history of cancer were used to identify families at risk for LS (Rodriguez-Bigas et al., 1997, Umar et al., 2004, Vasen et al., 1991, Vasen et al., 1999). Peltomäki *et al*. (Peltomäki et al., 1993) discovered the first susceptibility locus predisposing to LS, enabling genetic testing. MMR gene variants and their estimated pathogenicity classifications are collected in the InSIGHT DNA variant database (InSIGHT DNA Variant Database). Currently 320 Lynch syndrome families and more than 1700 diagnosed individuals have been registered in the National Lynch syndrome registry of Finland (LSRFi).

Identification of LS mutation carriers is important in order to offer access to regular surveillance programs such as colonoscopy and gynecological screenings to enhance the early detection of cancer and thus, increase the likelihood of survival. AC II and revised BG are the current guidelines used to identify suspected LS cases. Briefly, they highlight the family history of LS-associated cancer and diagnosis of LS-associated cancer under the age of 50 years (AC and revised BG). Additionally, AC require exclusion of Familial adenomatous polyposis (FAP) syndrome (Umar et al., 2004, Vasen et al., 1991). Revised BG specify the situations in which tumors should be tested for MSI when LS is suspected (Umar et al., 2004). In Finland, immunohistochemistry (IHC) is routinely performed on all CRC and endometrial cancer cases. With IHC, the likely gene(s) for subsequent mutational analysis can be identified (Kansikas et al., 2011, Yurgelun and Hampel, 2018). Additionally, MSI analysis is performed. Based on the IHC and MSI results, mutation screening is performed. Reliable sequencing for mutations detection of *PMS2* has proven demanding due to the 15 pseudogenes (Kasela et al., 2019).

5.2 Lynch-like syndrome (acquired condition mimicking Lynch syndrome)

Besides germline mutations in MMR genes (LS), inactivation of the MMR pathway may also be due to sporadic events: somatic methylation of the *MLH1* promoter, or biallelic

somatic changes such as somatic mutations or loss of heterozygosity (Geurts-Giele et al., 2014, Mensenkamp et al., 2014). Colorectal carcinomas that are MMR-deficient in the absence of *MLH1* promoter methylation or germline mutations (LS) represent Lynch-like syndrome (LLS) (Carethers, 2014). LLS patients typically fulfill AC II (Vasen et al., 1991) or revised BG (Umar et al., 2004) but lack molecular diagnosis of LS (germline mutations in MMR genes) (Carethers, 2014). However, not all LLS cases show family history of colorectal cancer or the family history may be unknown.

Improved molecular mechanisms in diagnosis and tumor testing have made it possible to identify LLS tumors in population-based cohorts. Only recently, Lynch-like syndrome (LLS) was established as a consistent group of MMR-deficient tumors by identification of double somatic mutations as the mechanism associated with the majority of the cases (approximately up to 70%). The double somatic events lead to silencing of both alleles of MMR gene enabling accumulation of replication errors in the genome (Geurts-Giele et al., 2014, Haraldsdottir et al., 2014). The mechanisms behind the rest, and the somatic and epigenetic profiles, are nevertheless, unknown. Earlier this year, Pearlman *et al*. (Pearlman et al., 2019) proposed to differentiate colorectal cancer cases with identified double somatic event from those cases with no obvious reason for MMR deficiency.

Comprehensive screenings for somatic mutations and epigenetic changes is expected to gain crucial information on LLS tumor-specific characteristics. Originally, term `Lynch-like syndrome` was used collectively to all MMR-deficient cases fulfilling AC II and/or revised BG with no obvious MMR germline mutation. In other words, out ruling the possibility of LS was the only means of LLS diagnosis. Recently, a more specific term “double somatic” has been suggested for cases in which the tumor is caused by two simultaneous mutations in a single MMR gene causing the deficiency (Pearlman et al., 2019).

AIMS OF THE STUDY

Germline mutations in MMR genes are well established as the causative factor in Lynch syndrome, but the somatic mutational and epigenetic profiles as well as the reasons associated with organ selectivity in the MMR mutation carriers remain to be unsolved. Only recently, LLS was classified as a separate group of MMR-deficient tumors by identifying double somatic mutations as the mechanism associated with the majority of the cases. The mechanisms behind the rest, and the somatic and epigenetic profiles, are nevertheless, unknown.

The specific aims were:

1. To determine the somatic mutation profiles and epigenetic status of the LS tumors of different organs, and to explore the interplay between genetic and epigenetic changes in LS tumorigenesis

2. To evaluate the mechanisms associated with MMR deficiency in a hospital-based cohort (patients operated in Jyväskylä Central hospital in 2000-2010), and the incidence of LLS in Finland

3. To explore whether germline mutations in MMR genes play a role in breast cancer tumorigenesis in LS mutation carriers

4. To investigate if LS synchronous ovarian and endometrial cancers have a shared origin (primary tumor and a metastatic lesion) or represent individually developing primary tumors

MATERIALS AND METHODS

6. Patient samples (I-IV)

The material of studies I, III and IV consist of tumor (LS colorectal carcinomas and adenomas, LS-CRCs and LS-ADs, LS ovarian carcinomas, LS-OCs, LS endometrium carcinomas, LS-ECs) and their counterpart normal samples (and hyperplasia samples in Study III) from individuals included in the National Lynch syndrome registry of Finland (LSRFi) that currently consists of approximately 300 families and more than 1600 tested LS carriers. In study III, we also identified all available breast carcinoma (BC) cases from the tested mutation negative family members of patients registered in LSRFi (non-carrier breast carcinomas, NC-BCs) that were collected as a control group. For the study II, 762 consecutive colorectal carcinomas diagnosed during years 2000-2010 in the Central hospital of Central Finland were collected. The number of tumors included in each of the studies are summarized in **Table 4** (see below 20.1. Patient characteristics: germline mutations and ages at onset). Please see the original publications I-IV for more information of the tumor characteristics. Our tumor and normal samples and patient information collections are unique sets of research material collected from LSRFi, relatives of patients registered to LSRFi, and from hospital region of Central Finland.

All tumor and hyperplasia samples and a proportion of normal samples (except for 31 blood samples in Study I) were formalin-fixed paraffin embedded (FFPE) tissue blocks which were collected from the archives. The histology of the tumors and tumor percentage of the FFPE samples was evaluated by a designated pathologist, and the samples were cut into 10 μm sections with a microtome for DNA extraction accordingly. Pathologist also evaluated a representative FFPE sample block and area in the tumor block of which the corresponding normal sample was cut and further extracted, for comparison.

Data analyzed in the studies is generated solely for each of the studies, and will not be reused in future studies. This novel data includes sequencing data, fragment analysis data, text files, and images, and it is stored appropriately (see 19. Ethical issues). Any data covered by copyright is not used. We require signed collaboration contracts with all the collaborators to ensure open science practice.

7. DNA extraction

DNA was extracted according to non-enzymatic protocols described by Isola *et al.* for FFPE samples (Isola et al., 1994) and Lahiri and Nurnberger for blood samples (Lahiri and Nurnberger, 1991).

8. Protein expression by immunohistochemistry (IHC) (I-IV)

Protein expression status of the tumor specimens was analyzed using immunohistochemistry (IHC) (**Table 2**). In studies I, II and IV, IHC was performed using 4 μm thick whole-slide (and tissue microarray (TMA) in study II) FFPE tissue sections, which were deparaffinized with xylene and dehydrated using graded alcohol. Counterstaining with hematoxylin (Mayers HTX, Histolab) was conducted after antigen retrieval, after which the tissue slides were cleared using xylene and mounted appropriately. In study III, IHC analysis was conducted using 2 μm TMA sections from FFPE blocks and following protocol: 15-minute incubation with 1mM EDTA/10mM Tris/HCI buffer in which the pH was 9.0 at 99 ^{O}C. Designated pathologists in each study reported their clinical findings of each sample.

Table 2. Summary of methods

Method	Study
Immunohistochemistry (IHC)	I-IV
Microsatellite instability (MSI) analysis	I-IV
Methylation-specific multiplex ligation-dependent probe amplification (MS-MLPA)	I-IV
4.5 Methylation-quantification of endonuclease-resistant DNA (MethyQESD) analysis	II
Exclusion of the Finnish founder mutation 1	II
Panel sequencing	I-IV
Somatic mutation analysis	I-IV
MMR second hit analysis	I-III
Characterization of the most mutated genes	I-III
In silico pathogenecity evaluation of somatic mutations	I
IonTorrent sequencing	II
Mutation signature analysis	III
Statistical analysis	I-IV

8.1 MMR-proteins (I-IV)

Protein expression of each the four MMR genes was measured using IHC staining. Undisputed nuclear staining in tumor cells demonstrates normal MLH1, MSH2, MHS6 and PMS2 expression, whereas negative immunostaining in cancer cells was interpreted to indicate inactivation of the MMR gene in question, combined with the internal positive control staining of the sample (blood vessels of stromal cells). The primary antibodies used are listed in **Table 3**.

Table 3. Summary of the primary antibodies used in IHC.

Study	Antibody	Clone	Dilution	Incubation time	Whole slide (WS)/ TMA	Tumor tissue	Manufacturer
I	Anti-MLH1	ES05	1:50	60 min	WS	Colon	Daco North America, USA
I	Anti-MSH2	G219-1129	1:1000	60 min	WS	Colon	BD Pharmingen, USA
I	Anti-MSH6	EP49, AC00-47	1:100	60 min	WS	Colon	Epitomics, USA
I	Anti-PMS2	EPR3947	1:1000	60 min	WS	Colon	Abcam, UK
I, IV	Anti-MLH1	G168-15	1:40	60 min	WS	Ovary, endometrium	BD Biosciences/ Pharmingen, Belgium
I, IV	Anti-MSH2	FE11	1:60	60 min	WS	Ovary, endometrium	Calbiochem/ Oncogene Research, Germany
I, IV	Anti-MSH6	44/MSH6	1:60	60 min	WS	Ovary, endometrium	BD Biosciences
I, IV	Anti-PMS2	A16-4	1:400	60 min	WS	Ovary, endometrium	BD Biosciences
II	Anti-MLH1	NCL-L-MLH1	1:100	60 min	TMA	Colon	Novocastra, Leica Biosystems, Germany
II	Anti-MSH2	NA27	1:150	60 min	TMA	Colon	Oncogene Research Products, USA
II	Anti-MSH6	287M-16	1:50	60 min	TMA	Colon	Cell Marque, USA
II	Anti-PMS2	556415	1:400	60 min	TMA	Colon	BD Pharmingen, USA
III	Anti-MLH1	G168-15	1:60	60 min	WS	Breast	BD Biosciences
III	Anti-MSH2	FE11	1:80	60 min	WS	Breast	Oncogene Research, Germany
III	Anti-MSH6	44/MSH6	1:40	60 min	WS	Breast	BD Biosciences
IV	Anti-L1CAM	1E11	1:40	20 min	WS	Ovary, endometrium	Covance
IV	Anti-ARID1A	HPA005456	1:200	20 min	WS	Ovary, endometrium	Sigma-Aldrich, USA

In study I, for visualization of MLH1, MSH2 and MSH6 antibodies, Ventana BenchMark XT immunostainer with OptiView detection system was used, whereas PMS2 was visualized using OptiView and Amplification detection system (Ventana Medical Systems, Tucson, AZ, USA). In Study II, visualization was performed using LabVision Autostainer 480 (Thermo Fisher Scientific, Fremont, CA, USA) and BrightVision + polymer detection kit (ImmunoLogic BV, Duiven, The Netherlands). In study III and IV, Dako EnVision+ (Dako, Glostrup, Denmark/Carpinteria, CA, USA) reagents were used according to manufacturer's protocol for visualization.

8.2 *BRAF* V600E (II)

BRAF^{V600E} hotspot mutation status was determined using IHC staining, a validated reliable method for *BRAF*^{V600E} mutation detection (Thiel et al., 2013). Mutation specific antibody (dilution 1:100, clone VE1, Spring Bioscience, Pleasonton, CA, USA) was used for immunostaining. A 25-minute incubation with 1 mM EDTA/10 mM Tris/HCl buffer with pH 8.0 at 99 ^{O}C was used for antigen retrieval. Positive staining was interpreted as indication of *BRAF*^{V600E} mutation whereas negative staining indicated wild-type *BRAF* status (Thiel et al., 2013).

8.3 ARID1A and L1CAM (IV)

For antigen retrieval, PT-Module (Lab Vision, Ca, USA) was carried out at 98^{O}C / 20 minutes in Evision TM Flex Target Retrieval solution. Solutions pH was 6.1 for ARID1A and 9.0 for L1CAM (Agilent Technologies, USA). For ARID1A we used rabbit produced antibody anti-ARID1A (20 minutes in dilution 1:200, HPA005456, polyclonal, Lot D104841, Sigma-Aldrich, USA), and mouse produced antibody Covance SIG-39110-200 for L1CAM (20 minutes in dilution 1:40, CD171, clone 1E11,

Covance). Staining was performed using Autostainer 480 automated immunostainer (Lab vision, CA, USA) and hematoxylin (Mayers HTX, Histolab) for counterstaining. When the nuclei of tumor cells lacked staining but the stromal cells stained positive, ARID1A expression was interpret as negative/abnormal. L1CAM expression was interpreted as positive/abnormal when > 10% of tumor cells showed L1CAM expression.

9. Microsatellite instability (MSI) analysis (I-IV)

Microsatellite instability (MSI) status of the tumor specimens was analyzed using polymerase chain reaction (PCR) and mononucleotide repeat markers *BAT25* and *BAT26*. These five markers are shown to be specific and sensitive indicators of MSI status (Esemuede et al., 2010, Loukola et al., 2001). Tumors stable for both MSI-markers were considered as microsatellite-stable (MSS), whereas tumors in which one or both repeat markers were unstable were classified as microsatellite unstable (MSI). Samples were classified as MSI-high when 2-5 of markers indicated MSI, MSI-low when only one of the five markers indicated MSI, and MSS when none of the five markers indicated MSI.

In study II, MSI analysis was conducted for part of the samples using IHC. Exhibition of positive staining for all four MMR genes was considered as an indication of MSS phenotype of the sample, whereas samples that exhibited undisputedly negative for one or more marker genes were considered as MSI based on Shia *et al*. 2008 (Shia, 2008).

The tumors were considered MMR-deficient (dMMR) if the IHC staining was abnormal, tumors exhibited MSI phenotype, or both. If both IHC staining was normal and tumor showed MSS phenotype, tumor was classified as MMR-proficient (pMMR).

10. Methylation analysis (I-IV)

The methylation data produced in the studies I-IV were primarily conducted using methylation-specific multiplex ligation-dependent probe amplification (MS-MLPA) method described by Nygren *et al.* (Nygren et al., 2005), except in study II, in which Methylation-quantification of endonuclease-resistant DNA (MethyQESD) method was the primary method used to define *MLH1* methylation status. We used 200-600 ng of DNA and followed the manufacturer's instructions in all MS-MLPA analyses (MRC Holland). Briefly, MS-MLPA analysis is based on probes containing recognition sequence (GCGC) that is recognized by methylation-sensitive endonuclease HHaI (Promega, Wisconsin, USA), which digests the GCGC site by binding to its unmethylated CpG dinucleotide. Digestion can only occur if the CpG dinucleotides within the restriction site are unmethylated because HhaI enzyme only recognizes unmethylated CpG dinucleotides. When the C in the CpG dinucleotide is methylated, sequences remain undigested and can be amplified generating a signal peak detectable in PCR. After PCR, the products were separated by capillary electrophoresis using ABI 3730 Automatic DNA sequencer (Applied Biosystems, Carlsbad, CA, USA). The MS-MLPA test produces a methylation dosage ratio (Dm), which varies between 0 and 1.0 illustrating the percentage of methylated DNA in the sample. Results were analyzed by genotyping software GeneMapper versions 4.0 and 5.0 (applied Biosystems, Carlsbad, CA, USA). For each tumor sample (and normal samples for comparison), the methylation dosage ratio (Dm) was calculated according to Pavicic *et al*. (Pavicic et al., 2011). The threshold for probe-specific hypermethylation in tumor tissues was determined against normal DNA: mean methylation Dm in normal samples plus two standard deviations. Conservative technical threshold for methylation detection was set to Dm value 0.15, which corresponds to 15% of methylated DNA (Nygren et al., 2005), and if the mean Dm plus two standard deviations was lower, Dm value 1.5 was used as cut-off.

10.1 CpG island methylator phenotype (I-III)

CpG island methylator phenotype (CIMP) phenotype is commonly used criteria in colorectal carcinomas to indicate the methylation status of tumor suppressor genes. Alternative 24 gene methylation panel was used for tumors originating from other tissues (see 10.2 Methylation of 24 genes). We used commercial SALSA MS-MLPA probemix ME042-B2 (MRC Holland, Amsterdam, The Netherlands) to study CIMP status of the colorectal specimens. CIMP panel covers the CpG islands in promoter regions of eight tumor suppressor genes *CACNA1G, IGF2*, *NEUROG1, RUNX3, SOCS1*, *CDKN2A*, *CRABP1*, and *MLH1*. CIMP probemix contains 3-5 probes per gene, and gene was considered methylated when ≥25% of the probes were methylated (Berg et al., 2014). Genes *CACNA1G*, *IGF2*, *NEUROG1*, *RUNX3*, and *SOCS1* were under our special interest as we followed the Weisenberger criteria in analyzing CIMP status (Weisenberger et al., 2006), according to which the sample is considered CIMP positive when at least three genes out of five (3/5) were methylated. Alternative Ogino panel including all eight CIMP genes where minimum five out of eight genes methylated (5/8) was used as comparison for CIMP criteria (Ogino et al., 2007).

As the CIMP panel also includes probes for *MLH1* promoter region, it was also used to determine the methylation status of *MLH1* in studies I and II. Furthermore, the CIMP panel includes a probe for sensitive and specific identification of $BRAF^{V600E}$ status, and this was utilized in some samples in study II.

10.2 Methylation of 24 genes (I-IV)

As the CIMP phenotype is commonly used and solely validated for colorectal tumors only, alternative methylation panel was used to define the methylator phenotype of tumors from other origins. MS-MLPA and SALSA MS-MLPA probemix ME001-C1 (in study III) and SALSA MS-MLPA probemix ME001-C2 (in study IV) (MRC Holland, Amsterdam, Netherlands) covering 24 general tumor suppressor genes was used. These

24 genes are methylated in several cancers, and specified in www.mrc-holland.com (MRC Holland). The panel contains 1 probe per gene, except for *RASSF1*, for which the panel contains two probes, and the gene was considered methylated if either of the two were methylated. The Dm value was calculated and 0.15 threshold was used for methylation, except for *CDKN2B* (in study IV, for which a Dm cut-off of 0.34 was set, due to higher than expected values reported by MRC-Holland), as follows: mean Dm in normal endometria was determined using LS normal endometrium samples, plus one standard deviation. Based on our previous experience, threshold of five or more genes methylated per sample was considered as hypermethylated, comparable to CIMP positive phenotype (Joensuu et al., 2008).

10.3 Custom panel (IV)

A custom MS-MLPA panel was designed, according to manufacturer's instructions, to study the methylation status of genes *RSK4*, *SPARC*, *PROM1*, *WT1-S*, *CABLES1*, *HOXA10*, and *HOXA9*, often methylated specifically in ovarian and endometrial cancer (MRC Holland). This custom design panel included one probe for 7 genes supplemented with SALSA MS-MLPA kit P-300-B1 human DNA reference-2 reagents. The Dm cut-off value was individually calculated for each gene the way it was described for *CDKN2B* gene above (see 10.2 Methylation of 24 genes).

10.4 *MLH1* methylation status (II)

Colorectal samples with deficient MLH1 expression detected by IHC were further studied for *MLH1* promoter methylation in study II. Methylation-quantification of endonuclease-resistant DNA (MethyQESD) was the primary method used. Quantitative MethyQESD method combines methylation-sensitive digestion with realtime PCR as described in (Bettstetter et al., 2008). Briefly, the methylation-sensitive endonuclease

Hin6I only recognizes and cuts unmethylated CGCG sites, whereas methylated samples remain uncut. This difference in proportions of methylated vs. unmethylated DNA can then be determined by real-time PCR. Methylation threshold of 16.6% or more was used as indication of methylation, as described (Bettstetter et al., 2008). Parallel *MLH1* methylation analysis by MS-MLPA was conducted as *MLH1* is included in the SALSA MSMLPA probemix ME042-B2 (see 10.1 CpG island methylator phenotype). MS-MLPA and other supportive data suggested, however, that some cases with methylation percentage of ≥11.7% were considered as methylated.

11. Finnish founder mutation-specific test (II)

Majority of LS patients in Finland carry a 3.5 kb genomic deletion of exon 16 and its flanking introns in *MLH1*. This specific mutation is called the 'Finnish founder mutation' due to its high incidence in the Finnish population. Due to the large size of the deleted region, this mutation cannot be detected by standard sequencing. Therefore, a mutation specific test was conducted for colorectal tumors with absent MLH1 expression using IHC in study II, in order to exclude the mutation in this Finnish population-based collection, as described by Nyström-Lahti *et al*. (Nystrom-Lahti et al., 1995).

12. Panel sequencing (I-IV)

Tumor and paired normal DNA samples were delivered to the Institute for Molecular Medicine Finland (FIMM) for panel sequencing. We chose to target known cancer genes by using the Nimblegen Comprehensive Cancer Panel (CCP) (Roche Diagnostics), which is a 4 Mb design targeting 578 cancer-related genes and their intronic regions compiled from the Sanger Institute Cancer Gene Census database (Sanger institute Cancer Gene Census) and the NCBI Gene tests database (NCBI Gene tests database).

ThruPLEX® DNA-seq Kit was used for library preparation and exon capture was conducted according to manufacturer's protocol (Rubicon Genomics), after which the sequencing was performed on Illumina HiSeq 2500 platform (San Diego, CA, USA). The variant calling pipeline (3.4 in study I and 3.6 in studies II-IV) is described in detail by Sulonen *et al.* (Sulonen et al., 2011). Briefly, raw Illumina reads were merged with SeqPrep, after which the resulting paired reads were trimmed of B blocks in the quality scores from the end of the read, from which only reads of 36 or more base pair long reads were selected for further processing. These reads were aligned against the human genome GRCh37 reference-genome primary assembly using the Burrows-Wheeler Aligner version 0.6.2 (Li and Durbin, 2009), and the alignment was refined using GATK Indel Realignment version 3.4. After the alignment, reads mapping to multiple genomic positions were removed, and potential PCR duplicates were removed using Picard MarkDuplicates version 1. 90.

12.1 Somatic mutation analysis (I-III)

Somatic mutation analysis was performed by FIMM to identify mutations of somatic origin using the data from panel sequencing of the tumor and paired normal samples. Non-synonymous somatic mutations (missense, nonsense, frameshift, in-frame coding deletion/insertion and splice site mutations) were selected for further analysis using VarScan 2 mutation detection algorithm version 2.3.2 (Koboldt et al., 2012). High-confidence somatic mutations were called using the following parameters: strand-filter 1, min-coverage-normal 8, mincoverage-tumor 6, somatic-p-value 1, normal-purity 1, and min-var-freq 0.05. SnpEff version 4.0 (Cingolani et al., 2012) and Ensembl v68 annotation database (Yates et al., 2016) were used for mutation annotation. Misclassified germline variants were filtered out by removing the common population variants included in the Database of Single Nucleotide Polymorphisms (Database of Single Nucleotide Polymorphisms (dbSNP), Build ID:130).

Only variants with VarScan somatic p-value below 0.01 were considered as significant and true, and only those were selected for subsequent analyses in studies I-IV. For this reason, the term `somatic mutation` is referred to any non-synonymous sequence change with the possibility of being pathogenic (including traditional pathogenicity classes 3 - 5) throughout this book.

13. Ion Torrent sequencing (II)

Simultaneously with the panel sequencing conducted in FIMM, 15 of the colorectal tumors in study II were sequenced with Ion Torrent in Central hospital of Central Finland in Jyväskylä, Finland to verify the panel sequencing findings. *MLH1, MSH2, MSH6* and *PMS2* coding and untranslated regions were covered by Ion Ampliseq™ (Thermo Fisher Scientific) custom panel, which is designed for 125–175 bp amplicon range comprising 163 amplicons that covers 92.36% of the target areas. Standard manufacturer's protocol (Ion Ampliseq™) was used in library preparation, and sequencing was conducted using Ion torrent PGM (Thermo Fisher Scientific). Torrent Suite™ Software (Thermo Fisher Scientific) was used for primary data processing, and Ion reporter version 4.2 (Thermo Fisher Scientific) for variant calling. Variants with less than 20× coverage were filtered out.

14. MMR second hit analysis (I-III)

Second hit analysis was performed to identify the second, somatic mechanism leading to MMR gene deficiency. Comprehensive Cancer panel (CCP) Sequencing data of MMR genes was studied to identify possible somatic second hit point mutations, and the promoter methylation of the MMR genes was studied by MS-MLPA in each sample.

To identify the possibility of loss of heterozygosity (LOH) as a second hit, LOH analysis was performed. When the primary mutation was a point mutation, VarSeq software (GoldenHelix®) with VCP filtered panel sequencing data (.vcf-files) was used to compare the data from tumor and their corresponding normal samples. The variant allele reads (Alt) vs. reference allele reads (Ref) ratio was determined in both tumor (T) and normal (N) DNA, and LOH ratio (R) calculated as follows: $R = (Alt:Ref)_T/(Alt:Ref)_N$. The thresholds for LOH and putative LOH specified by Ollikainen *et al.* were used (Ollikainen et al., 2005). When the primary mutation was a large deletion, MLPA produced data (SALSA P003-C1 for *MLH1* and *MSH2* and SALSA 072-C1 for *MSH6*, MRC Holland, Amsterdam, The Netherlands) was used in LOH analysis and the results interpreted according to Zhang *et al*. (Zehir et al., 2017). Putative and strict LOH are called LOH throughout this thesis.

15. *In silico* analysis of somatic mutations (I-III)

VarSeq (GoldenHelix®) was used for *in silico* evaluation of the somatic single nucleotides variants (SNVs) in studies I, II, and IV. We used six individual algorithms included in VarSeq to predict the effect of amino acid substitution on protein function: SIFT (SIFT), PolyPhen-2 (Adzhubei et al., 2010), MutationTaster (Schwarz et al., 2014), MutationAssessor (MutationAssessor), FATHMM (Shihab et al., 2013a, Shihab et al., 2013b, Shihab et al., 2014), and FATHMM MKL Coding (FATHMM).

Pathogenicity classifications of MMR somatic mutations were done by checking against the InSIGHT database (InSIGHT database (Leiden Open Variation Database, LOVD v. 2.0 Build 36) in studies I, II, and IV. Furthermore, somatic mutations in MMR genes and somatic mutations in the top mutated genes (see 16. Identification of the top-mutated genes, below) were also assessed against the Catalogue of somatic mutations in cancer (Catalogue of somatic mutations in cancer (COSMIC)), and Human Splicing Finder

(Human Splicing Finder) was used to predict the splicing consequences of SNVs in splice site regions in the second hit analysis (see 14. MMR second hit analysis).

In study III, *in silico* predictions to assign a category of pathogenic significance of somatic mutations in MMR genes were performed using Varsome-database (Varsome - The Human Genomics Community).

16. Identification of the potential driver genes (top-genes) (I-III)

To identify the genes with the highest rate of mutations, the potential driver genes, we developed a method used in studies I-III. First, the proportion of tumors in which a particular gene was mutant was calculated for each of the 578 genes. Second, focusing on high allele frequency mutations (frequency >25%) to increase the likelihood of clonal (driver) as opposed to subclonal (passenger) mutations (Williams et al., 2016), and to the distribution of mutated genes in LS colorectal tumors in Study I, a cut-off of 31% was established to divide the tumors into groups of commonly and less commonly mutated. This cut-off was subsequently applied to LS ovarian carcinomas in study I, LLS tumors in study II, and breast carcinomas in Study III, to enable tumor type-specific comparisons. Third, for each of the identified top-genes, a pathway annotation was specified according to GeneCards (GeneCards, Human Gene Database) and relevant publications from PubMed (PubMed). Finally, top-mutated genes were compared to the list of 719 genes in Cancer Gene Census v85 (Cancer Gene Census).

17. Mutational signature analysis

Non-synonymous somatic mutations identified by the VarScan2 mutation detection algorithm version 2.3.2 (see 12.1. Somatic mutation analysis) were selected to mutation

signature analysis conducted by R package deconstructSigs (Rosenthal et al., 2016) against the signatures recognized by the COSMIC database (Catalogue of somatic mutations in cancer (COSMIC)). The following parameters were used: normal-purity 1, strand-filter 1, min coverage 8 and 6 (for normal and tumour samples, respectively), minimum variant frequency 0.05, and somatic p-value 1. The deconstructSigs applies multiple linear regression model to the input data, thus, determining the individual mutational profiles of the tumor samples. In the analysis, individual signatures of each sample were combined with signatures of other tumors in each group (dMMR LS-BSc, pMMR LS-BCs, NC-BCs, LS-CRCs, LS-OCs, and LLS-CRCs), and average of each signature was calculated. Signatures with frequency lower than 5% were combined to group 'Other'. Validation of breast cancer signatures was conducted by identifying all somatic high-confidence variants by Mutect2 (MuTect2) and analyzing the signatures by MutationalPatterns R-package, according to Blokzijl *et al* (Blokzijl et al., 2018).

18. Statistical analyses

The statistical analyses presented in the studies I, II, and IV and in this thesis book were performed using IBM SPSS Statistics software (IBM SPSS Inc., Chicago, IL, USA) version 24 and version 25 in study III. The data's applicability for parametric or non-parametric tests was first tested. The statistical significance of mutations or mutated genes' distribution in independent groups was evaluated using the Mann-Withney U test. The pairwise comparisons of frequency data was conducted by the Fisher's exact test. The Pearson correlation coefficient for parametric data was used to assess the correlations. Two-tailed p-values < 0.05 determined the significant values.

19. Ethical issues

Studies I-III are approved by the Institutional Review Board of the Helsinki University Central Hospital (466/E6/01). In addition, study II is approved by the ethics committee of Jyväskylä Central Hospital (Dnro 13U/2011). Study IV was approved by the Institutional Review Boards of the Departments of Surgery (466/E6/01) and the Obstetrics and Gynecology (040/95) of the Helsinki University Central Hospital (Helsinki, Finland) and Jyväskylä Central Hospital (Jyväskylä, Finland) (Dnro 5/2007). The National Supervisory Authority for Welfare and Health (Dnro 1272/04/044/07 and Dnro 10741/06.01.03.01/2015) approved the use of the patient registry and collection of archival specimens. Written consent forms were always obtained at the time of sample collection.

Sensitive data is handled according to the EU's general data protection regulations, and it is never shared outside our research group. An important part of confidentiality consists of handling the patient samples and clinicopathological information as unique codes.

RESULTS

20. Overview of the clinicopathological characteristics of the patients and tumors (Studies I-IV)

20.1. Patient characteristics: germline mutations and ages at onset

The majority of Lynch syndrome patients carried germline mutations in *MLH1*, accounting for 39/57 (68%) of the tumors. This reflected the prevalent Finnish founder mutation I, a 3.5 kb genomic deletion of exon 16 and its flanking introns in *MLH1*. Moreover, *MSH2* and *MSH6* germline mutations were identified in 9/57 (16%) carriers each. (**Table 4**). None of the tumors were from germline mutation carriers of *PMS2*.

The mean age of onset was 44, 46, 58, and 49 years for LS-CRCs, LS-OCs, LS-BCs (dMMR and pMMR combined), and LS-ECs, respectively. The mean age of onset for all the LS tumors combined was 49 years. Among non-LS tumors studied, the mean age of onset was 59 years for NC-BCs and 65 years in LLS-CRC, and 62 years combined (**Table 4**).

Table 4. Summary of clinicopathological characteristics of patients and panel sequenced tumors

		Study	Average age of onset	Mean no. of non-synonymous somatic mutations	Proportion hypermutated [> 10 (ns) mutations per Mb]	Predisposing gene			Proportion of hypermethylation phenotype (CIMP)	Two-hit inactivation			
						MLH1	*MSH2*	*MSH6*		Germline mutation + LOH	Germline mutation + somatic point mutation(s)	No obvious second hit	ND
LS CRC	dMMR[#] (n=18)	I-III	44	689 (172/Mb)	18/18 (100%)	12/18 (67%)	2/18 (11%)	4/18 (22%)	9/18 (50%)	9/18 (50%)[¤]	6/18 (33%)[¤]	2/18 (11%)[¤]	1/18 (6%)[¤]
LS-OC	dMMR[#] (n=16)	I-IV	46	735 (184/Mb)	13/16 (81%)	13/16 (81%)	3/16 (19%)	0	3/15 (20%)	7/16 (44%)[¤]	3/16 (19%)[¤]	1/16 (6%)[¤]	5/16 (31%)[¤]
LS-BC	dMMR[#] (n=11)	III	53	696 (174/Mb)	10/11 (91%)	7/11 (64%)	2/11 (18%)	2/11 (22%)	0[$]	6/11 (55%)	4/11 (36%)	1/11 (9%)	0
	pMMR[#] (n=9)	III	63	131 (33/Mb)	4/9 (44%)	4/9 (44%)	2/9 (22%)	3/9 (33%)	0[$]	2/9 (22%)	3/9 (33%)	3/9 (33%)	1/9 (11%)*
LS-EC	dMMR[#] (n=3)	IV	49	583 (146/Mb)	3/3 (100%)	3/3 (100%)	0	0	1/3 (33%)	DM			
LLS-CRC	dMMR[#] (n=14)	II	65	124 (31/Mb)	11/14 (79%)	N/A			13/14 /93%)	N/A			
NC-BC	pMMR[#] (n=10)	III	59	352 (88/Mb)	3/10 (30%)	N/A			0[$]	N/A			

N/A, Not applicable

DM, Data missing

[#] Based on MMR deficiency detected by IHC and/or MSI positive status

[$] Based on data in Lotsari et al. 2012

[¤] Based on data in Porkka et al. 2017

* Second hit analysis was inconclusive in one tumor because LOH analysis failed

20.2 Comparison of somatic mutation counts and CIMP statuses between tumor groups (Studies I-IV)

Panel sequencing of 578 cancer-relevant genes revealed the average number of 689 (172/Mb), 735 (184/Mb), and 414 (103/Mb) non-synonymous somatic mutations per LS-CRCs, LS-OCs, and LS-BCs, respectively. When observing LS-BCs separately by their MMR-status, an average of 696 (174/Mb) somatic mutations in dMMR vs. 131 (33/Mb) in pMMR tumors were identified (p=0.015) (**Table 4**). Majority of LS tumors were hypermutated: proportion of hypermutated tumors was 100% (18/18), 81% (13/16), 70% (14/20), and 100% (3/3) in LS-CRCs, LS-OCs, LS-BCs (dMMR and pMMR combined), and LS-ECs, respectively (**Table 4**).

The combined group of LS tumors (LS-CRC, LS-OC, and LS-BC both dMMR and pMMR; n=54) revealed significantly more somatic mutations (average 611) compared to breast carcinomas from non-carrier individuals (NC-BCs, n=10) (average 352 somatic mutations, p=0.032) and LLS colorectal carcinomas (n=14) (average 124 somatic mutations, p=0.027). Furthermore, the number of mutated genes out of the 578 CCP panel genes was significantly higher in LS tumors combined (average 190 somatic mutations), compared to NC-BCs (average 119, p=0.038) and borderline significant compared to LLS-CRCs (average 82, p=0.050).

Hypermethylation (CIMP) was detected in 50% of LS-CRCs, 20% of LS-OCs, and 33% of LS-ECs. No hypermethylation phenotype was observed in LS-BCs. Among non-LS groups, hypermethylation phenotype was absent in NC-BC, but particularly frequent in LLS-CRC (93%) (**Table 4**).

20.3 Somatic mutation counts and CIMP statuses in Lynch-like syndrome tumors (Study II)

In this thesis and in the study II, we refer to LLS when talking about tumors that are MMR-deficient with no identified germline mutations in MMR genes or *MLH1* promotor methylation. This includes tumors with identified double somatic changes in MMR genes, tumors with only one identified somatic mutation in MMR genes, as well as tumors in which no obvious causative mechanism could be identified in our study. The average age of CRC onset in LLS patients was 65 years, which was significantly higher compared to LS-CRCs (44 years, p=0.001), significantly lower compared to CRCs with hypermethylated *MLH1* (76 years, p=0.001), and lower than sporadic pMMR CRCs from Study II (70 years, p=not significant).

Panel sequencing of LLS-CRCs identified an average of 124 (31/Mb) somatic mutations per tumor, which was significantly lower compared to LS-CRCs (689, p<0,001) or all studied LS-tumors combined (average 611, p=0.027) (**Table 4**).

Unlike the LS tumors, most LLS-CRCs proved to be CIMP positive (13/14, 93%) without *MLH1* promoter hypermethylation (**Table 4**). To the best of our knowledge, our study was the first to report the association between LLS-CRCs and CIMP positive phenotype.

20.4 Origin of synchronous Lynch syndrome ovarian and endometrial tumors (Study IV)

The study of synchronous ovarian and endometrial carcinomas from LS-carriers was conducted to resolve weather the tumors are individually arising primary tumors or metastatic lesions of one tumor. The genetic and epigenetic characteristics of these tumors indicate a shared origin, that is, the tumors are metastatic lesions of one primary tumor in two locations. This result is concordant with sporadic synchronous ovarian and

endometrial carcinomas that are also clonally related representing metastatic disease (Anglesio et al., 2016, Schultheis et al., 2016).

All 13 synchronous carcinoma pairs were molecularly concordant based on MMR status, ARID1A expression, L1CAM expression, and hypermethylation status for 7 and 24 genes (see Materials and Methods). Enough DNA material was available for targeted sequencing of five tumor pairs, of which three pairs showed undoubtedly shared origin with identical high confidence somatic mutations. A fourth tumor pair shared lower-confidence somatic mutations (detected with relaxed criteria, accepting somatic *p*-value <0.05), and was hence interpreted as of possibly shared origin. The fifth tumor pair only shared a single somatic mutation with relaxed criteria, leaving the origin unsettled.

The complex endometrial hyperplasias and the endpoint lesion endometrial or ovarian carcinomas exhibited high molecular concordance, suggesting convergence of the ovarian and endometrial tumorigenic pathways associated with the metastatic disease.

20.5 Mechanisms of two-hit inactivation for MMR genes (Studies I-III)

In LS tumors (LS-CRCs, LS-OCs, and LS-BCs), where the germline MMR gene mutation constitutes the first hit, somatic inactivation of the second allele of the deficient MMR gene was investigated (**Table 4**), whereas in LLS-CRCs somatic inactivation of both first and second allele was studied. The main mechanism of second hit in the remaining active allele in LS-tumors was LOH (44%, 24/54). In 30% of the LS-tumors (16/54) the second hit was a somatic point mutation. The second hit in the remaining 35% of LS-tumors remained to be discovered, partly due to insufficient DNA material (**Figure 5**). By the techniques we used, the allele specificity could not be determined.

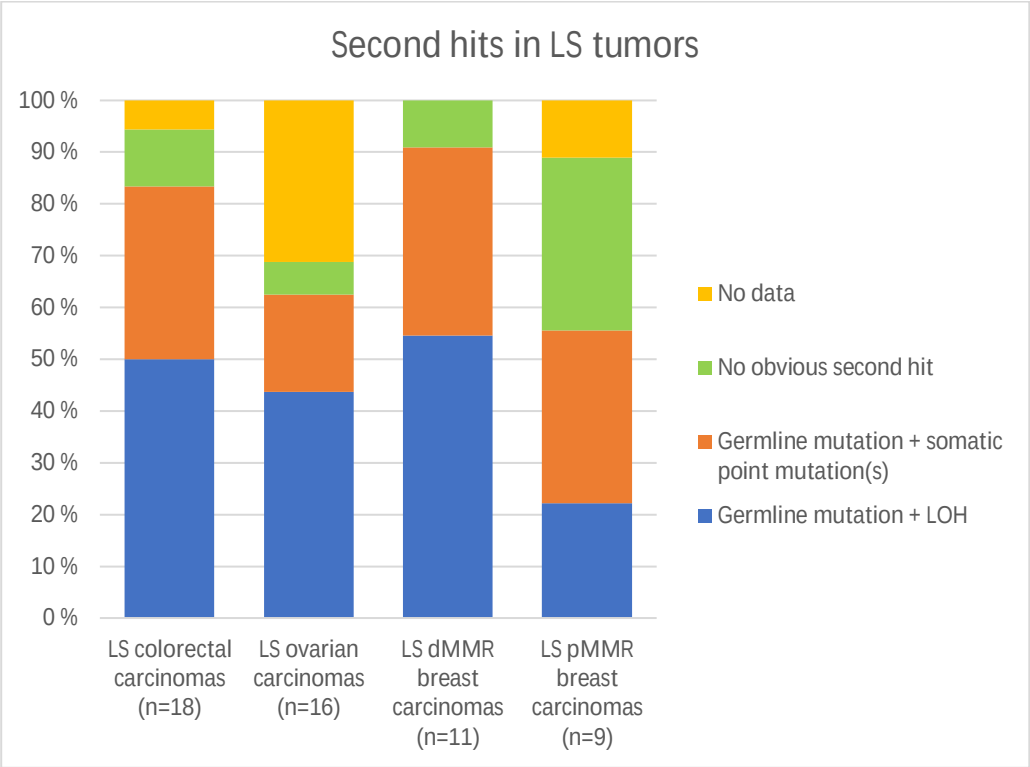


Figure 5. First and second hit in LS tumor groups. Based on second hit analysis, LOH was the main mechanism of the second hit in majority of LS tumors, and the main second hit in all MMR-deficient tumors (LS CRC, LS ovarian carcinomas and LS dMMR breast carcinomas). The second hit was more often detected in the dMMR tumor groups, provided that enough material was available for the analysis (insufficient DNA material was available in 5/16 of LS ovarian carcinomas).

Somatic MMR gene point mutation together with loss of heterozygosity provided the major mechanism accounting for LLS-CRCs (**Figure 6**). Interestingly, double somatic point mutations in MMR genes were identified in 3/10 (30%) of NC-BCs, although they exhibited pMMR phenotype.

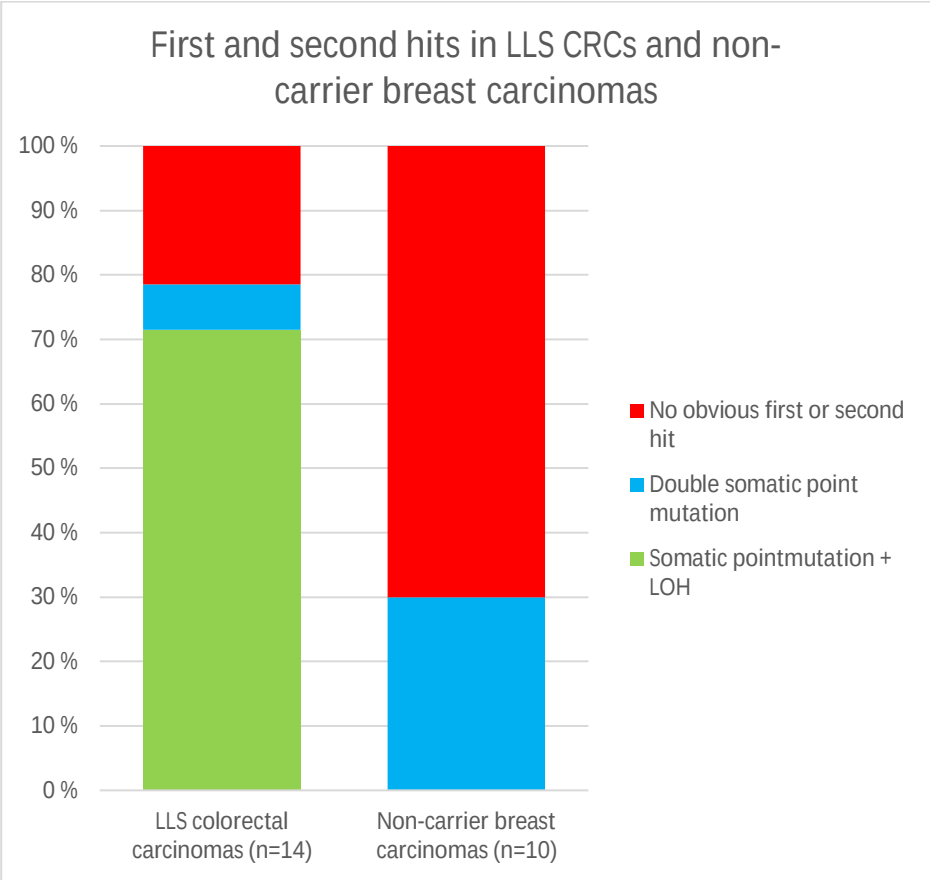


Figure 6. First and second hits in LLS colorectal carcinomas and non-carrier breast carcinomas. The first and second hit was identified in the majority of LLS CRCs (79%), compatible with the tumors being MMR-deficient. In non-carrier breast carcinomas (all MMR proficient), double somatic events in MMR genes were present in 30%.

21. Unique features for genetic and epigenetic profiles of different tumors (Studies I-III)

21.1 Higher somatic mutation burden for hypermethylated tumors (Studies I-IV)

MMR status was an important classifier of somatic mutation burden. When classifying the tumors into MMR-deficient (dMMR, n=59) (LS-CRCs, LS-OCs, dMMR LS-BCs,

and LLS-CRCs) and MMR proficient (pMMR, n=19) (pMMR LS-BCs and NC-BCs), dMMR tumors had significantly more somatic mutations (average 569) compared to pMMR tumors (average 247, p=0.004). Also, the number of genes mutated out of 578 genes was significantly higher in dMMR tumors (average 181 vs. 101, respectively, p=0.004). LS-EC was not considered due to low number of samples.

Apart from the MMR status, tendency to acquire somatic mutations depended on the presence vs. absence of a generalized hypermethylation phenotype (CIMP). When observing the somatic mutational counts in all hypermethylated (n=25) and non-hypermethylated (n=52) tumors (LS-CRC, LS-OC, LS-BC, NC-BC, and LLS-CRC), the average somatic mutation counts and number of mutated genes (/578) were significantly higher in hypermethylated tumors (702 vs. 378, p=0.046, and 207 vs. 143, p=0.034, respectively). When considering only the LS tumors (hypermethylated n=12, non-hypermethylated n=41), the difference was even more significant: average 1318 vs. 419 somatic mutations, p=0.00045, and 337 vs. 152 somatic mutations, p=0.00049, respectively. Average numbers of somatic mutations in different tumor groups are shown in **Figure 7**. The three endometrial carcinomas sequenced in study IV did not comply with the general pattern since one that was hypermethylated had 132 somatic mutations, whereas the two non-hypermethylated tumors had 72 and 1544 mutations.

Finally, a difference was observed between hypermethylated tumors from inherited vs. acquired MMR gene mutation carriers: hypermethylated LS-CRCs (9/18) and LS-OCs (3/16) had significantly more somatic mutations (843 and 2745, respectively) compared to LLS-CRCs, of which 13/14 were hypermethylated (124, p<0.001 and p=0.003, respectively).

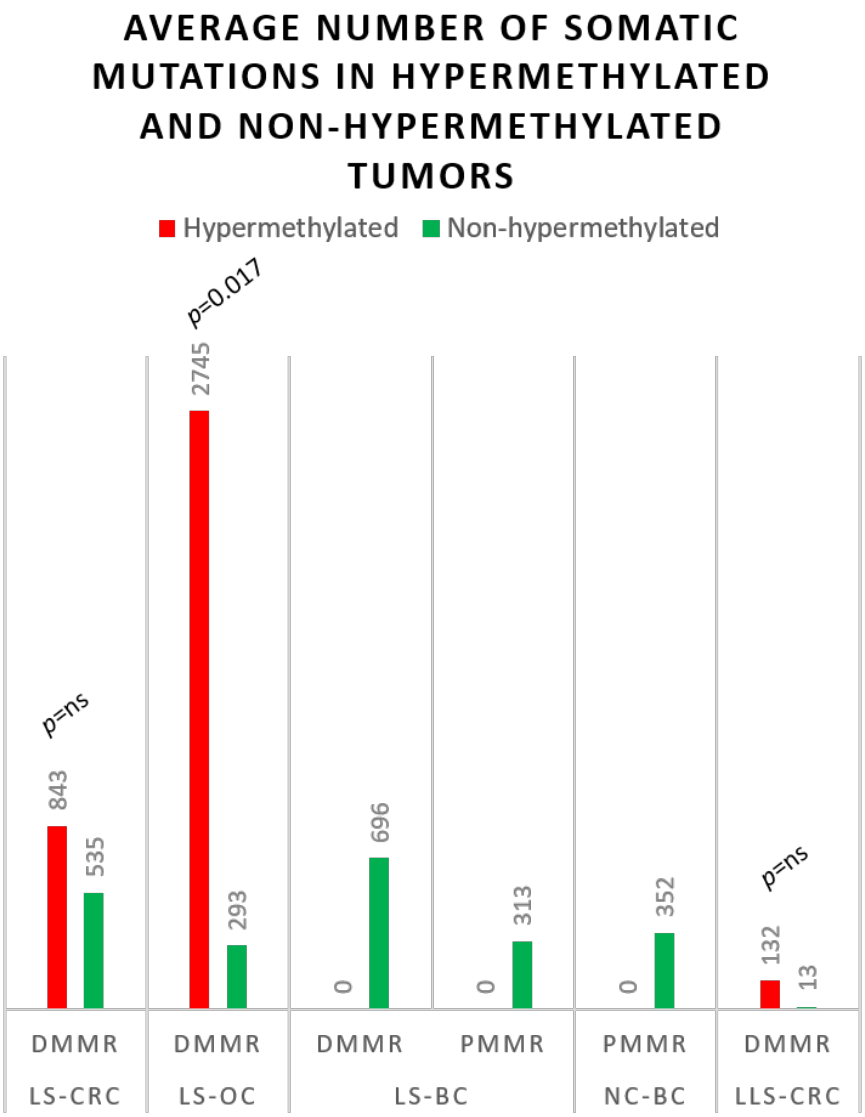
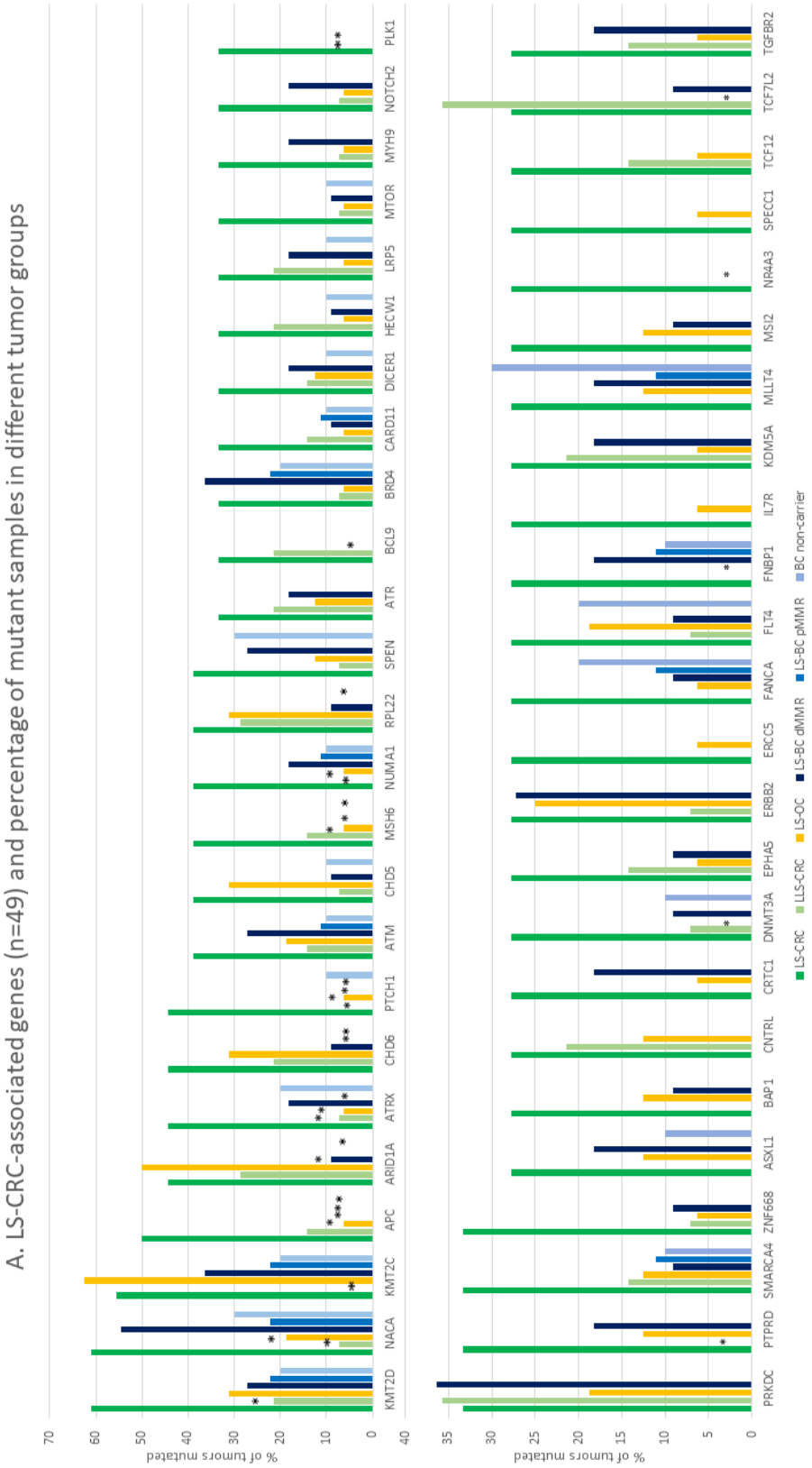


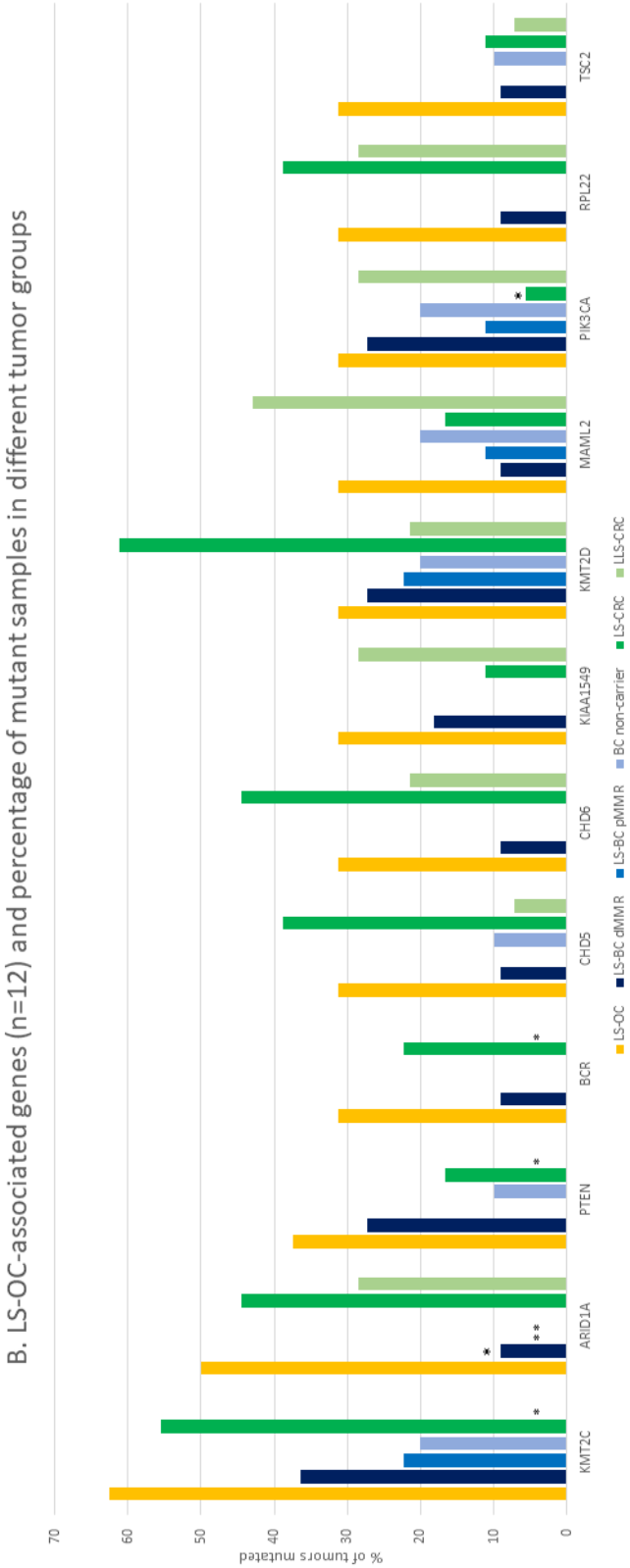
Figure 7. Average number of somatic mutations in hypermethylated and non-hypermethylated tumor groups. Comparison of somatic mutational counts in tumors with hypermethylation (red bars) and no hypermethylation (green bars) revealed more somatic mutations in hypermethylated tumors. Significant difference was observed in ovarian carcinomas (LS-OC) (p=0.017). Endometrial carcinomas are not presented in the figure due to the low number of samples (n=3). Data concerning the BC samples is from Lotsari et al. (Lotsari et al., 2012).

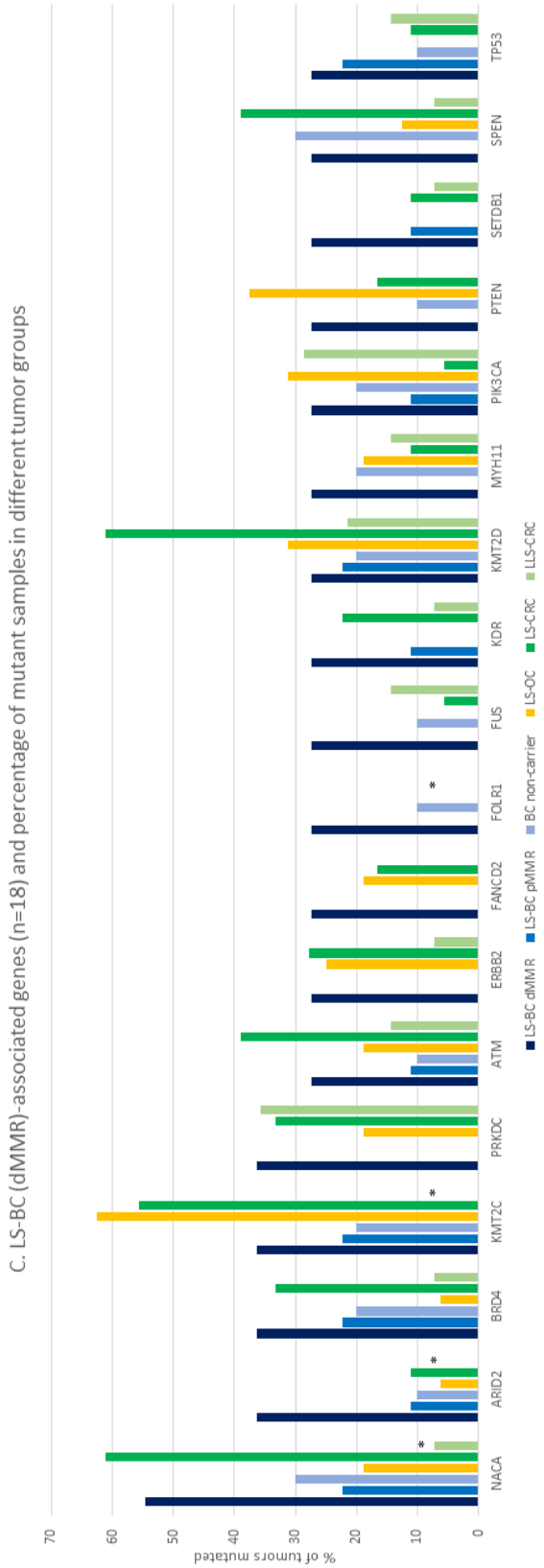
21.2 Unique somatic mutation profiles for tumors from different organs (Studies I-III)

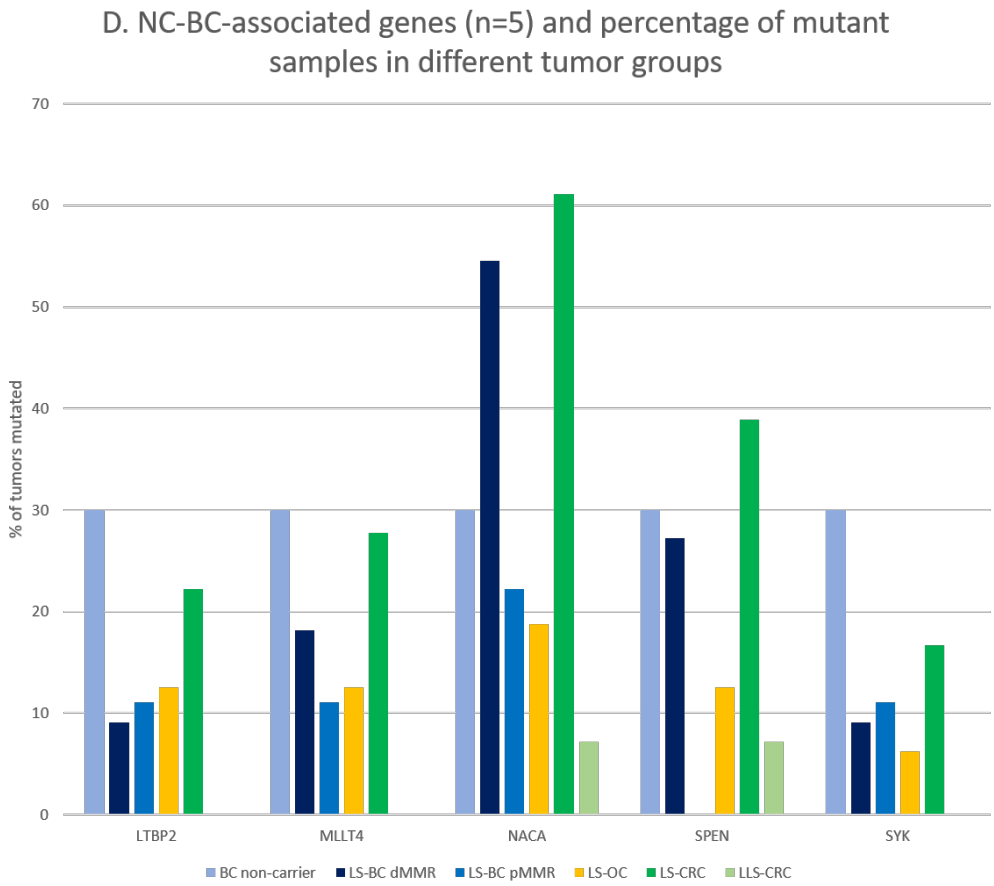
Somatic mutation profiles of different LS tumors have remained poorly characterized. Panel sequencing of 578 cancer related genes revealed that LS tumors bear unique somatic mutation profiles. Genes harboring mutations with allele frequency of 25% or higher in approximately 30% of tumors were regarded to be possible drivers. Among 49, 12, and 18 such "top mutated" genes identified for LS-CRCs, LS-OCs, and dMMR LS-BCs, respectively, the most frequently mutated genes were *KMT2D* and *NACA* (61% of LS-CRCs), *KMT2C* (63% of LS-OCs), and *NACA* (55% of dMMR LS-BCs) (Figure X A – C (**Figure 8**). In pMMR LS-BCs, genes with mutations in only 22% of samples at maximum were observed; hence, no driver genes meeting our selection criteria were identified.

LLS-CRCs also had a unique set of top mutated genes compared to LS-CRCs (**Figure 8 E**). The most mutated gene was *MAML2* that was mutated in 43% of the LLS tumors. Only four out of 13 of the top genes were shared with LS-CRCs: *PRKDC*, *TCF7L2*, *ARID1A*, and *RPL22* were mutant in 33%, 28%, 44%, and 39% of LS-CRCs, respectively. *BCOR* was among the top mutated genes of LLS-CRCs, and it was not mutated with any of the LS-CRCs (p=0.028).


A. LS-CRC-associated genes (n=49) and percentage of mutant samples in different tumor groups
% of tumors mutated
KMT2D NACA KMT2C APC ARID1A ATRX CHD6 PTCH1 ATM CHD5 MSH6 NUMA1 RPL22 SPEN ATR BCL9 BRD4 CARD11 DICER1 HECW1 LRP5 MTOR MYH9 NOTCH2 PLK1
PRKDC PTPRD SMARCA4 ZNF668 ASXL1 BAP1 CNTRL CRTC1 DNMT3A EPHA5 ERBB2 ERCC5 FANCA FLT4 FNBP1 IL7R KDM5A MLLT4 MSI2 NR4A3 SPECC1 TCF12 TCF7L2 TGFBR2
LS-CRC LLS-CRC LS-OC LS-BC dMMR LS-BC pMMR BC non-carrier


B. LS-OC-associated genes (n=12) and percentage of mutant samples in different tumor groups
% of tumors mutated
0
10
20
30
40
50
60
70
KMT2C
ARID1A
PTEN
BCR
CHD5
CHD6
KIAA1549
KMT2D
MAML2
PIK3CA
RPL22
TSC2
LS-OC
LS-BC dMMR
LS-BC pMMR
BC non-carrier
LS-CRC
LLS-CRC


C. LS-BC (dMMR)-associated genes (n=18) and percentage of mutant samples in different tumor groups
% of tumors mutated
0
10
20
30
40
50
60
70
NACA
ARID2
BRD4
KMT2C
PRKDC
ATM
ERBB2
FANCD2
FOLR1
FUS
KDR
KMT2D
MYH11
PIK3CA
PTEN
SETDB1
SPEN
TP53
LS-BC dMMR
LS-BC pMMR
BC non-carrier
LS-OC
LS-CRC
LLS-CRC


D. NC-BC-associated genes (n=5) and percentage of mutant samples in different tumor groups
% of tumors mutated
70
60
50
40
30
20
10
0
LTBP2
MLLT4
NACA
SPEN
SYK
BC non-carrier
LS-BC dMMR
LS-BC pMMR
LS-OC
LS-CRC
LLS-CRC

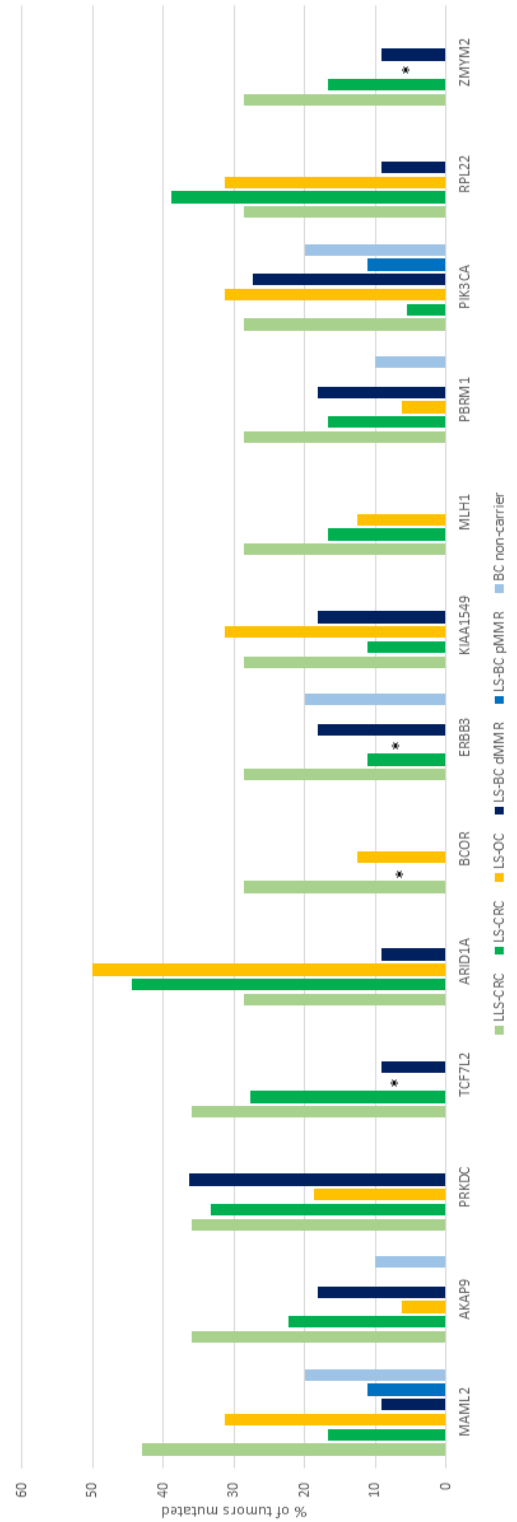


Figure 8 A-E. Percentage of mutated tumors for top mutant genes characteristic of each tumor group. Shares of mutated tumors for top-mutated genes characteristic of each tumor group are presented in Figures A-E. The group of tumors whose top genes are shown and to which other tumor groups are compared is depicted as the first bar in the figure for each gene. Significant differences in shares between tumor groups are identified by asterisk (*). The sample sizes of the tumor groups were 18 for LS-CRC, 16 for OC, 11 for dMMR LS-BC, 9 for pMMR LS-BC, 10 for NC-BC, and 14 for LLS-CRC.

22. Molecular classification of 762 colorectal carcinomas from a population-based cohort (Study II)

In the stepwise study of population-based cohort of 762 CRCs, 655 pMMR (86%) and 107 dMMR (14%) CRCs were identified by immunohistochemical analysis of tumors for MMR protein expression. The dMMR tumors were investigated further for *MLH1* promoter methylation and germline mutations of MMR genes, resulting in 81 *MLH1* hypermethylated (11%) and 9 LS cases (1%). The remaining 14 tumors (2%) with neither *MLH1* methylation nor LS germline mutations were classified as LLS tumors (**Figure 9**). LS accounted for 8% and LLS 13% of dMMR CRCs.

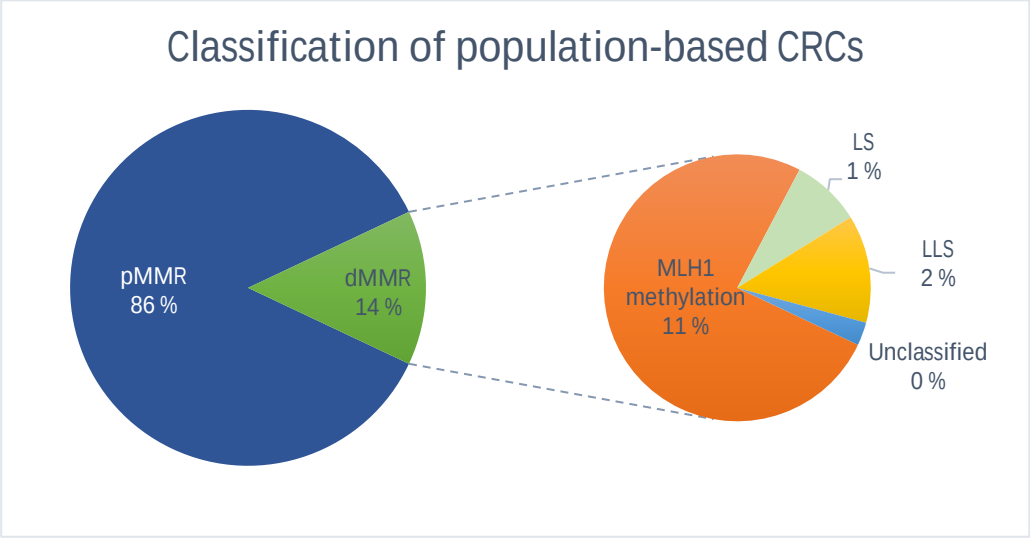


Figure 9. Molecular classification of 762 colorectal carcinomas from a population-based cohort. A step-wise characterization of 762 population-based colorectal carcinomas identified that 86% of the tumors were MMR proficient (dark blue) and 14% MMR-deficient (green), of which further analysis enabled identification of *MLH1* promotor hypermethylated (11%, orange), Lynch syndrome (1%, light green), and Lynch-like syndrome (2%, yellow) tumors. Three tumors remained unclassified (<1%, light blue).

23. Mutational signatures of LS tumors (Study III)

Mutation signature analysis in study III was performed to detect the individual signatures of LS-BCs and compare them to the signatures of breast carcinomas from non-carrier family members and LS tumors originating in other organs (**Figures 10 and 11**). Signatures of individual samples were first analyzed. For each signature, average over sample-specific values representing a specific tumor group was determined and used to evaluate differences between the tumor groups. Signatures of dMMR LS-BCs resembled other LS tumors more closely than signatures of NC-BCs. MMR deficiency-associated signature 6 was present (frequency >5%) in all LS tumors but not in NC-BCs. Other MMR deficiency associated signatures 20 and 26 were also found in dMMR LS-BCs, LS-CRCs (only signature 26), and LS-OCs with >5% frequency (see details in Study III). Notably, in addition to signature 6, pMMR LS-BCs also exhibited signature 20 with frequency of 14%. NC-BCs also exhibited signatures 6 and 20 with low frequencies <5%. When counting all the MMR deficiency associated signatures 6, 20, and 26 combined, NC-BCs had the lowest frequency (5%) and pMMR LS-BCs the highest frequency (23%) (for comparison, the corresponding frequencies were 12% in dMMR LS-BCs, 13% in LS-CRCs, and 16% in LS-OCs).

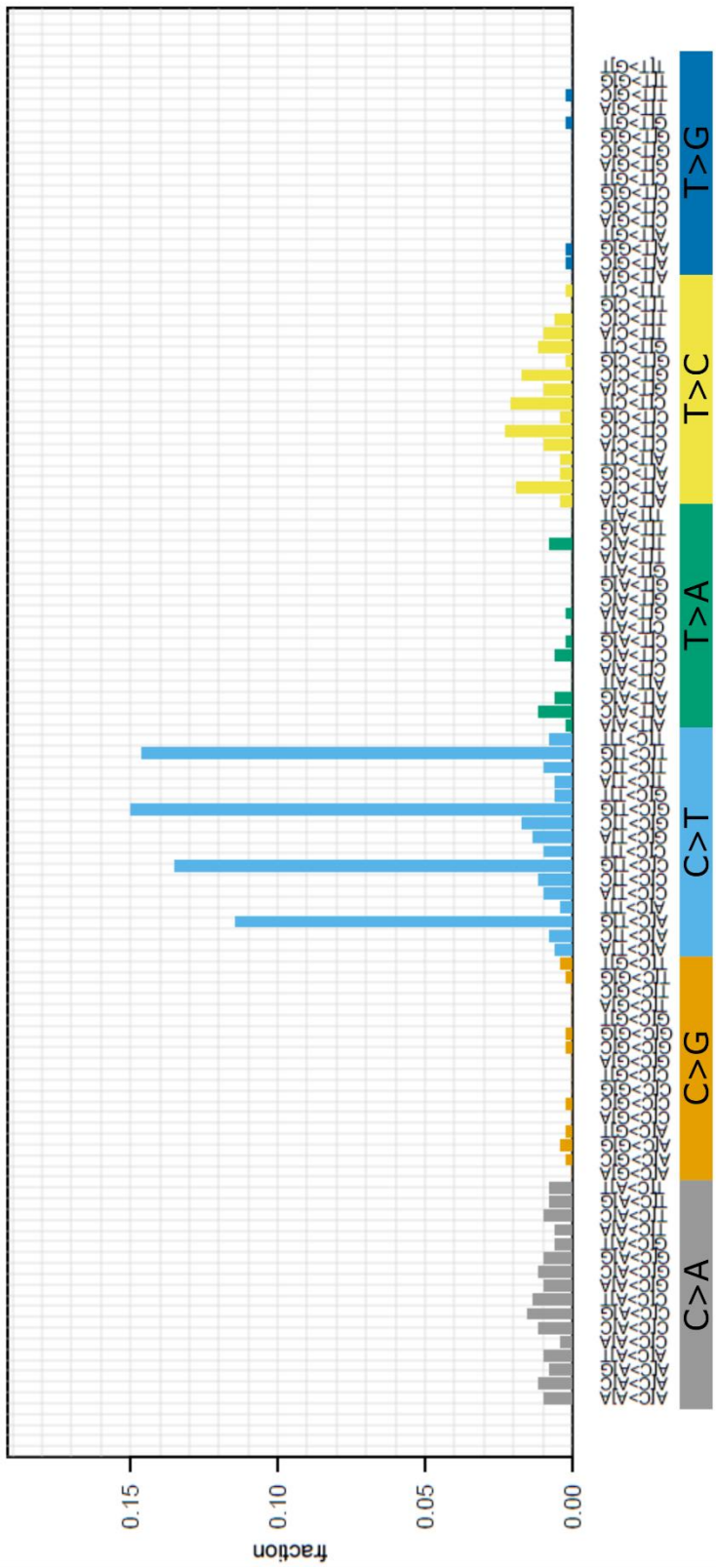


Figure 10. Mutation signature of an individual dMMR LS-BC sample by the deconstructSigs. The deconstructSigs determines the individual mutational profiles of each tumor samples by applying multiple linear regression model to the input data. Signatures 1 and 6 are present in this particular sample with frequencies of 85% and 15%, respectively. Signature 6 is associated with MMR deficient phenotype and MSI tumors (COSMIC).

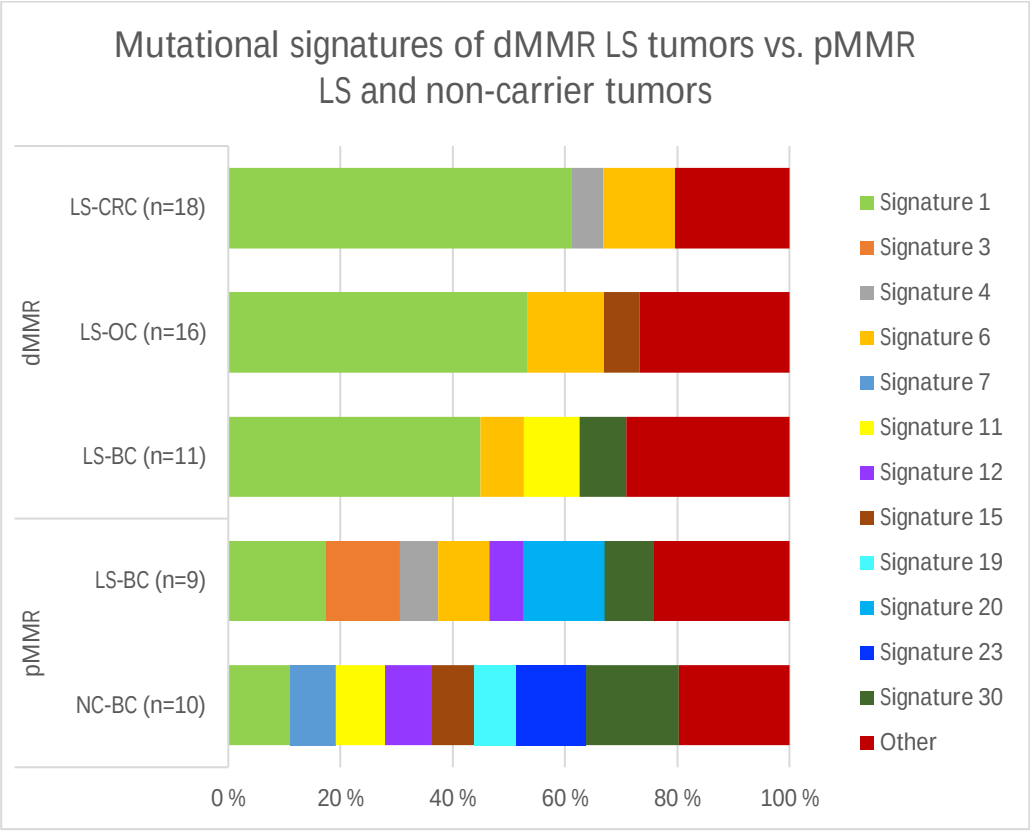


Figure 11. Mutational signatures of different LS tumor groups. Mutational signatures of LS tumor groups and non-carrier breast cancer group classified by their MMR-status. Tumors with dMMR had more similar mutational signatures compared with signatures of pMMR tumors. Signatures were calculated based on the averages of sample-specific signatures in each group. Signature 26 in dMMR BS-BCs, LS-OCs, and LS-CRCs are included in the red group 'other', as are all the other signatures with under 5% fraction.

DISCUSSION

24. Differential diagnostics of Lynch syndrome

24.1 Sporadic dMMR *MLH1* methylated CRC, LS CRC, and LLS CRC

Approximately 16% of all colorectal carcinomas show hypermutated phenotype, of which majority also exhibit high-degree MSI, which is a consequence of MMR deficiency (Cancer Genome Atlas, 2012). The majority, 80%, of MSI cases are sporadic, in which the mechanism associated with MMR deficiency is biallelic *MLH1* promoter hypermethylation silencing the *MLH1* gene (Boland and Goel, 2010, Lynch and de la Chapelle, 2003). Often, positive CIMP status and *BRAF* V600E mutation coexist with *MLH1* hypermethylated sporadic CRC (Lynch et al., 2007). Approximately 11% (81/762) of CRCs in our population-based study (Study II) showed *MLH1* hypermethylation (**Figure 9**). Majority (76%) of them carried *BRAF* V600E mutation in contrast to LS and LLS CRCs that were all negative for the mutation, and pMMR CRCs that were 95% negative for the mutation. Unlike in *MLH1* methylated CRCs, CIMP positive status did not correlate with *MLH1* hypermethylation in LS and LLS CRCs.

This thesis contributes to the increasing amount of information on MMR-deficient colorectal cancers beyond Lynch syndrome. Approximately 12-15% of CRCs are caused by deficiencies in MMR-machinery (Aaltonen et al., 1998, Moreira et al., 2012). Inherited germline mutations in MMR genes (Lynch syndrome) account for up to 15% of CRCs with deficient MMR expression or MSI phenotype, and 1-3% of unselected colorectal cancers (Peltomäki, 2014).

Lynch-like syndrome (LLS) forms another group of MMR-deficient CRCs exhibiting MSI phenotype (Carethers, 2014, Rodriguez-Soler et al., 2013). LLS tumors were initially classified as dMMR tumors with no promoter hypermethylation of *MLH1*

(Geurts-Giele et al., 2014, Haraldsdottir et al., 2014, Mensenkamp et al., 2014), and therefore the only means to diagnose LLS was to exclude LS. Rodriguez-Soler *et al*. were the first to classify MMR-deficient tumors with no germline mutations or promoter methylation of *MLH1* as LLS, suspecting that the causative MMR mutation was not found due to methodological reasons or the complexity of the mutation. In other words, LLS cases were assumed to be a heterogeneous group consisting of sporadic CRCs combined with LS cases in which the suspected germline mutation in MMR genes could not be identified (Rodriguez-Soler et al., 2013). They found that the standardized incidence ratio of CRC was the lowest in families with pMMR or sporadic tumors (0.48), and highest in LS families (6.04), whereas the incidence ratio for LLS families (2.12) was in between, suggesting increased familial cancer risk in LLS.

Recently, another method for classification has been proposed. Pearlman *et al*. suggested that since the incidence of double somatic mutations as the causative factor in many of the LLS cases is now known, the tumors harboring biallelic somatic mutations (pathogenic variant, PV) in MMR genes should be classified as double somatic (DS), distinguishing them from the unclassified cases, in order to highlight the underlying mechanism associated with MMR deficiency (Pearlman et al., 2019).

Our population-based study including 762 colorectal carcinomas identified 107 (14%) dMMR cases, of which 81 (76%) were proven to be caused by *MLH1* hypermethylation, 11 (10%) by double somatic events in MMR genes, and nine (8%) harbored germline mutation in MMR genes (LS). Furthermore, three cases (3%) showed MMR deficiency with no obvious reason to explain it, and three (3%) dMMR cases remained unclassified due to inadequate analyses of *MLH1* methylation or panel sequencing to determine the status. In our study (Study II), DS and unexplained dMMR CRC cases were classified as LLS (13%) (**Figure 9)**. None of these LLS cases harbored *BRAF* V600E mutation known to be associated with sporadic MSI tumors, consistent with previous studies (Cohen et al., 2016, Xicola et al., 2019).

Our results indicate that somatic MMR mutations together with loss of heterozygosity is the major mechanism accounting for LLS tumorigenesis. Of the non-hypermethylated dMMR tumors in our study, slightly higher proportion of DS cases were identified (11/26, 42%) compared to the cohorts in Perlman *et al*. (33% and 31%), and slightly lower proportion of LS cases (9/26, 35%) compared to the cohorts in Perlman *et al*. (58% and 45%) (Pearlman et al., 2019). The limited size of our cohort (762 in total, 26 non-hypermethylated dMMR CRCs) compared to the two cohorts in Perlman *et al*. (3346 in total investigated, 232 non-hypermethylated dMMR CRCs; 1182 in total, 51 non-hypermethylated dMMR CRCs) may likely cause this difference. Furthermore, we considered all mutations that had the possibility of being pathogenic (not known to be benign or likely benign) when observing somatic mutations in LLS cases, whereas Pearlman *et al.* only considered pathogenic variants as possible hits (Pearlman et al., 2019). Cohen *et al*. discovered significant increase in *PIK3CA* mutations in DS CRCs (67%), and less in LS CRCs (22%), *MLH1* hypermethylated (20%), and MSS (15%). *PIK3CA* was among the top-mutated genes identified in our study (4/14, 29%), although the difference was not significant compared to LS CRCs (1/18, 6% p=0.142).

The average age of CRC onset in our cohort was lowest in LS cases (44 years) compared to LLS cases (65 years, p=0.001), and highest in patients with *MLH1* hypermethylated tumors (76 years, p=0.001 compared to LLS). Pearlman *et al*. also discovered significantly higher age at diagnosis in LLS patients compared to LS patients: 59 (Ohio) and 69 (Iceland) LLS (DS) vs. 52 (Ohio) and 62 (Iceland) LS (p<0.001), respectively, but younger than estimated in sporadic tumors (Pearlman et al., 2019). Other studies have reported similar ages at diagnosis for LS and LLS: median ages of 48 years in LS vs. 53 years in LLS (Mas-Moya et al., 2015) and median ages of 49 years in LS vs. 55 years in LLS (Rodriguez-Soler et al., 2013), both significantly lower than median ages reported for sporadic (*MLH1* hypermethylated or mutant *BRAF* V600E) cancers (71 and 78 years, respectively).

Xicola *et al*. found a significant enrichment of mutations in DNA repair genes other than MMR genes that could possible explain the mutator phenotype in LLS tumors (Xicola

et al., 2019). There were 45 DNA repair genes in our panel including MMR genes. No significant enrichment was discovered in our cohort when observing the “top-mutated” genes required by our criteria, where a gene was required to be mutated in approximately 30% of the samples. Only *PRKDC* (36%) and *MLH1* (29%) of the 45 DNA repair genes were among the top mutated genes of LLS-CRCs, but they were frequently mutated in other LS-tumors too: 35% and 17% in LS-CRCs, 19% and 13% in LS-OCs, and 36% and 0% in dMMR LS-BCs, respectively.

Jansen *et al*. discovered that occasionally LLS (or suspected Lynch Syndrome, sLS) can be explained by germline or somatic *POLE/POLD1* variants together with consequent somatic MMR mutations (Jansen et al., 2016). The targeted CCP panel, however, did not include *POLD1* or *POLE* genes and hence the mutation status of these genes in our LLS tumor cohort remains unknown.

LLS and DS cases are relatively new categories awaiting comprehensive screening for somatic mutations and epigenetic changes and comparison with LS tumors, such as we performed, to gain more information on LLS tumor-specific characteristics. Identification of their unique characteristics will enable improved diagnostics to separate them from LS tumors, and development of targeted treatment specific to molecular features of the tumors. Our finding of regular CIMP-positive phenotype (not affecting promoter region of *MLH1*) and somatic mutational profiles characteristic for LLS CRCs may facilitate diagnostics and guide the development of targeted treatments.

24.2 LS breast and ovarian cancer vs. *BRCA1/2*-associated and sporadic cases

Breast and ovarian cancer share rare genetic mechanisms, such as mutations in *BRCA1* and *BRCA*2, as well as environmental exposures associated with endogenous and exogenous hormone exposures (Hulka, 1997). Furthermore, a recently published large

genome-wide association study by Jiang *et al.* reported for the first time a significant genetic correlation between breast and ovarian cancers (Jiang et al., 2019).

Tumor suppressor genes *BRCA1* and *BRCA*2 are involved in homologous recombination repair. They are the most common susceptibility genes for hereditary breast and ovarian cancer with mutations causing predisposition to Hereditary Breast and Ovarian Cancer Syndrome (HBOC) (Futreal et al., 1994, Wooster et al., 1995). Inherited predisposition causes 5-10% of all breast cancers, and HBOC accounts for approximately 20-25% of this predisposition (Futreal et al., 1994, Wooster et al., 1995). Majority (80%) of pathogenic mutations in *BRCA1* and *BRCA2* are hereditary, that is, germline mutations (Riaz et al., 2017). The cumulative lifetime risk in women carrying germline mutation in either of these genes is significantly increased compared to general population: 65% in *BRCA1* and 45% in *BRCA2* mutation carriers for breast cancer, occurring at a younger age, and 39% in *BRCA1* and 11% in *BRCA2* mutation carriers for ovarian cancer (Juvet and Natvig Norderhaug, 2008). It is important to diagnose the predisposing mutation since mutations in *BRCA1* and *BRCA2* are highly penetrant and often associated with aggressive forms of breast or ovarian cancer. A significant fraction of *BRCA1* -associated breast cancers do not express estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor (HER-2), and are hence considered as triple-negative. In contrast, breast carcinomas from *BRCA2* mutation carriers are rarely triple-negative and hence resembled sporadic tumors in this respect (Atchley et al., 2008, Mavaddat et al., 2012). *BRCA1/2*-associated ovarian cancer is typically of high-grade serous histological type (McLaughlin et al., 2013), which also predominates among sporadic ovarian carcinomas (Kurman and Shih Ie, 2008).

After HBOC, the second most common cause of hereditary ovarian cancers is Lynch syndrome, accounting for approximately 2% of all ovarian cancers, and 8-13% of hereditary ovarian cancers (Malander et al., 2006, Walsh et al., 2011). Unlike ovarian carcinomas caused by HBOC, LS ovarian carcinomas generally represent well-differentiated early stage tumors with mainly non-serous histology (Helder-Woolderink et al., 2016).

Studies published to date have contradictory results on whether breast cancer should be regarded as a part of LS tumor spectrum (Aarnio et al., 1999, Buerki et al., 2012, Lotsari et al., 2012, Pande et al., 2012, Saita et al., 2018, Watson et al., 2008, Win et al., 2012). A large body of literature reporting molecular characterization of the LS breast tumors and patients have yielded strong evidence that LS breast cancer should be considered as LS-related cancer (Engel et al., 2012, Lotsari et al., 2012, Roberts et al., 2018, Saita et al., 2018, Win et al., 2012). We have previously reported that breast carcinomas from LS carriers resembles sporadic breast carcinoma in many clinicopathological respects, but LS carcinomas with respect to genomic instability (Lotsari et al., 2012). In Study III we discovered that dMMR LS-BCs exhibited mutational signatures comparable to signatures in other LS-tumors and different from signatures from non-carrier BCs (sporadic). Furthermore, both dMMR and pMMR LS-BCs showed signatures associated with MMR-deficiency (**Figure 11**).

Multiple studies suggests that LS is caused by mutations primarily in *MLH1* and *MSH2* genes causing carriers to have the highest risk for colorectal, endometrial, and ovarian cancers (Bonadona et al., 2011, Hampel et al., 2006, Hampel et al., 2008, Moller et al., 2017a, Moreira et al., 2012, Palomaki et al., 2009). The cohorts in these studies, however, often represent individuals with personal and/or family history of colorectal and/or endometrial cancer, making the cohorts less ideal for the evaluation of the risk for breast cancer in LS (Espenschied et al., 2017). Moreover, several studies have shown that germline mutations in *MSH6* and *PMS2* are associated with breast cancer in LS (Desmond et al., 2015, Espenschied et al., 2017, Minion et al., 2015, Roberts et al., 2018, Tung et al., 2016). However, the study of Roberts *et al*. reporting 2-3 fold increased risk for breast cancer in *MSH6* and *PMS2* germline carriers has been criticized for not reporting the reason for which the genetic testing was originally conducted (Roberts et al., 2018). Previous reports from the same research group indicate that more than half would have been tested because of diagnosis of breast cancer, and the cohort would therefore be misrepresentative. Recently, Dominguez-Valentin *et al*. published a large multicenter prospective study reporting the breast cancer risk to be comparable across

all four MMR genes, with only a minor increase in the cumulative risk to ages 60 and 75 (7-8% and 12-15%, respectively) compared to the risk in general population (Dominguez-Valentin et al., 2019). Other prospective studies have reported similar results with no increased risk for breast cancer in LS mutation carriers compared to general population (Moller et al., 2017a, Moller et al., 2017b, Moller et al., 2018). Our cohort of breast cancers from LS mutation carriers harbored germline mutations in *MLH1* (11/20, 55%), *MSH2* (4/20, 20%), and *MSH6* (5/20, 25%). The reasons for large proportion of *MLH1* carriers and non-existent *PMS2* carriers can be explained by the Finnish founder mutation in *MLH1* and by the rarity of *PMS2* mutations in Finnish population. Guidelines set by the National Comprehensive Cancer Network (NCCN) state that solely the "suggestions" of increased breast cancer risk in LS patients are not enough to endorse "increased screening above-average-risk breast cancer screening recommendations" for LS carriers (Sorscher, 2019).

A study by Espenschied *et al.* (Espenschied et al., 2017) described that 27.3% (158/579) of the MMR mutation carriers in their cohort did not meet any current criteria (Amsterdam criteria or Bethesda guidelines) for LS testing. Furthermore, *BRCA1/2* (and not LS) testing criteria were significantly more likely met in *MSH6* and *PMS2* carriers than in *MLH1* and *MSH2* carriers. They also reported that the available tumor MSI/IHC results were discordant with the mutations in 12.5% of *MHS6* and 18.2% *PMS2* gene associated cases, which complies with previous findings about lower IHC/MSI sensitivity in *MSH6* and *PMS2* cases (Terdiman, 2005). Thus, *MSH6*- and *PMS2*-associated cases are more likely to be missed by current LS testing (MSI/IHC), possibly making them underrepresented in previous LS studies (Espenschied et al., 2017) as well as in our studies.

While the average age of onset in BC was similar in LS-BCs (57 years) and NC-BCs (59 years) in our study, it was lower in dMMR LS-BCs compared to pMMR LS-BCs (53 years versus 63 years, p=0.036). This is higher than reported by Vasen *et al.* (46 years on average) and Roberts *et al.* (50 years on average) (Roberts et al., 2018, Vasen et al., 2001). The mean age of onset in LS-OCs in Studies I-III was 46 years, which is

concordant with published reports (45 years), and up to 20 years younger than reported in sporadic ovarian carcinomas (Helder-Woolderink et al., 2016).

Characterization of the molecular profiles of tumors is important for diagnosis and treatment. The finding that metastatic MMR-deficient breast carcinomas responded extremely well to immune checkpoint treatments, as reported by Kok *et al.* (Kok et al., 2017), provides an illustrative example. Approximately half (55%) of the LS-BCs in Study III were dMMR, which is concordant with previous studies (Win et al., 2013). Based on the tissue of origin, MMR and MSI status may vary in LS carriers due to the growth pattern of MMR deficient cells which is characterized by clonal heterogeneity (Peltomäki, 2017). All the NC-BCs were pMMR, concordant with the reports stating that sporadic dMMR breast cancers are extremely rare (Davies et al., 2017). The molecular characteristics between LS-BCs and sporadic NC-BCs as well as between dMMR and pMMR LS-BCs differentiated the groups from each other. In addition to the mutational signatures, each BC group had their unique profiles of somatic mutations (**Figure 8**).

The molecular profiles of synchronous endometrial and ovarian carcinomas (Study IV) indicate shared origins of the tumors, that is, metastatic disease, consistent with synchronous sporadic cases (Schultheis et al., 2016). However, conclusions of direction of the metastasis cannot be made based on our data. Kelemen *et al*. discovered that ovarian carcinomas synchronous with endometrial carcinoma highly resembled non-synchronous endometrial carcinomas and to a lesser degree resembled non-synchronous endometrioid ovarian carcinomas, implying metastasis from the endometrium to the ovary (Kelemen et al., 2017). Another study suggests that endometrioid and clearcell ovarian carcinomas have metastasized from endometrium (Karnezis et al., 2017). Our data showed the strongest molecular resemblance between the synchronous ovarian and endometrial carcinomas, and close molecular similarity between synchronous ovarian or endometrial carcinomas and the non-synchronous cancers of the same tissue of origin. However, endometrial hyperplasias were relatively frequent in patients with only

ovarian carcinomas, which might indicate a continuum from endometrial precusors to ovarian cancer.

25. Methods used in genetics and epigenetics

The past two decades have been extremely important in the field of human genetics. The Human Genome Project conducted the first sequencing of the entire human genome in 2001 (International Human Genome Sequencing, 2004, Lander et al., 2001, Venter et al., 2001). The sequencing method used in this project was traditional Sanger sequencing which was extremely time-consuming and took approximately 15 years (Green et al., 2015). In the following decade, commercial sequencing methods were developed allowing the emergence of Next Generation Sequencing (NGS) methods that have revolutionized cancer genetics. The ever-evolving and faster NGS methods have enabled more cost effective sequencing of targeted regions, whole-exome, and whole-genome, decreasing the costs from billions to thousands of dollars in just over two decades (Morganti et al., 2019). Databases, such as The Genome Aggregation Database (gnomAD) including 120 000 exomes and 15 000 whole genomes (Genome Aggregation Database, gnomAD) are available for reference and control purposes.

NGS methods have enabled increasing and more detailed information on hereditary and somatically acquired illnesses, such as cancer and many others. Based on this information, more efficient diagnostic methods as well as targeted therapies have been developed. Standard stratification of tumor mutation profiles and tailored management of patients based on tumors mutational signatures is the future goal in oncological care (Morganti et al., 2019).

The function of all genes in the human genome are not known, and with today's knowledge, the function of only a fraction of the entire genomic area is known. Moreover, variant classification poses a challenge and the 'interpretive gap' leads to a

large proportion of the variations in the genome being classified as 'variants of unknown significance' (VUS) because their effect in gene function is unknown (Beroud et al., 2016, Cutting, 2014). This raises a limitation to exome and whole-genome sequencing results, because a lot of information with unknown significance is produced, as well as lot of possibly significant information about other traits not associated to the original disease. Is it necessary to produce all the data if it cannot be utilized or which could possibly reveal information that has a significant impact on one's health but is not associated with the original condition that motivated the sequencing? Then again, the information about human genome is constantly increasing, making it possible to reinterpret the results in the future.

Large genomic deletions and insertions as well as copy-number variations (CNVs) are known to occur in many cancers. NGS data may be used for detecting large genomic rearrangements, such as CNVs (Schmidt et al., 2017). In studies I-IV, we used a mutation specific test to identify the large deletion in *MLH1* that is the Finnish founder mutation, but other large deletions and insertions, complex rearrangements, as well as CNVs would not have been reliably detected by the used panel testing method targeting the coding regions. Deletions in *EPCAM*, secondarily inactivating *MSH2* by hypermethylation leading to MMR-deficiency (Kuiper et al., 2011) would likewise have been missed in our studies. We also did not study changes in microRNAs (miRNAs), which are important regulators of many cellular processes (Peng and Croce, 2016). miRNAs are small non-coding RNAs that are associated with regulation of gene expression and have been found to be highly dysregulated in cancer cells (Peng and Croce, 2016)

We chose to use a commercial gene panel covering 578 known cancer-related genes and their intronic regions to limit the amount of information with unknown significance. To identify true positive variants, somatic mutation analysis was conducted to detect the non-synonymous somatic mutations. This, however, may affect for example the mutational signature analysis since only non-synonymous variants in targeted regions of the genome are covered, although reports that panel sequencing and/or restriction to non-synonymous variations has been used to capture successful signatures do exist (Nowak

et al., 2017, Zehir et al., 2017). There is a possibility that relevant mutations are located in genes not included in the CCP-panel, which must be taken into consideration when interpreting these results. Furthermore, majority or the LS-tumors in our data sets were hypermutated limiting the effect of restriction to non-synonymous mutations. However, all the samples in our studies were processed alike, so these limitations only concerns comparisons to signatures reported by others. In the interest to detect clonal changes, we focused mainly on somatic mutations with high allele frequency (≥ 25%); however, this does not take into consideration the possibility of somatic mosaicism in which *de novo* variants may exist in low frequencies (Renaux-Petel et al., 2018). Consequently, the somatic mutations in MMR genes in study II would have been missed if they only occurred in a small population of cancer cells. Published studies focusing on the somatic mutation spectra of LS or LLS tumors are scarce, limiting the possibilities for comparisons between investigations.

CIMP analysis with defined indicator markers is an established method to study the methylation status of CRCs, but no established criteria exists for other tumor types. Based on our previous studies (Niskakoski et al., 2013), the methylation status of tumors other than CRCs was determined by methylation analysis of 24 tumor suppressor genes. The methods to study the gene methylation have developed rapidly during the past two decades, enabling genome-wide methylation analysis (Choukrallah et al., 2019). In the studies I-IV, these genome-wide methods, however, were not utilized, in part because of FFPE origin of our specimens. In the epigenetic studies, we only focused on the promotor methylation analysis of CIMP panel genes, panel of 24 tumor suppressor genes, and promoter methylation of *MLH1* (the latter as the cause of deficient expression of MLH1), and this limitation should be taken into consideration when interpreting the methylation results.

26. Future aspects

In the future, larger cohorts of samples preferably representing multiple populations would be needed to verify our preliminary results on LS and LLS tumors. The epigenetic changes together with hypermutability, and especially their cause and effect relationships should be further studied. The predictive value of positive CIMP status remains to be clarified in LLS patients, taken into consideration its association with worse prognosis of colorectal cancer irrespective of MSI status. Furthermore, LLS tumors potential response to PD-1 blockade-based immunotherapy should be studied, as significantly better response has been observed in non-Lynch patients with MMR-deficient tumors. Comparison with corresponding data derived from sporadic tumors could also be relevant; for example, it would be interesting to study whether the correlation we observed between CIMP/hypermethylated phenotype and high somatic mutation burden also applies to sporadic tumors and which mechanisms are involved. The molecular differences between dMMR and pMMR breast cancers, especially in LS patients, should also be taken into consideration when considering the best treatments as well as when new treatments are developed.

SUMMARY AND CONCLUSIONS

NGS and epigenetic methods enable comprehensive characterization of somatic mutations and methylation profiling of different tumors. Detailed characterization of LS tumors from different tissues as well as LLS tumors bring important new information that may be used for diagnosis and targeted treatment. Many of the somatically mutated genes belong to signaling pathways to which targeted treatments exists or could be developed.

Study I

- LS tumors harbored organ-specific profiles of somatic mutations
- Hypermethylated (CIMP positive) LS colorectal and ovarian tumors gained more somatic mutations in cancer-relevant genes compared to non-hypermethylated (CIMP negative) tumors
- Genes involved in epigenetic regulation were enriched among the top mutated genes (potential driver genes)
- Signaling pathways such as DNA repair, Wnt, mTOR, PI3K, NOTCH, MAPK, and tyrosine kinase were affected. These pathways may potentially be targeted by for example pan-HER inhibitors in *ERBB2*-mutant MSI colorectal cancer

Study II

- LLS encompassed a significant proportion of MMR-deficient colorectal cancers (17%), and most LLS tumors (11/14, 79%) were due to double somatic mutations
- LLS colorectal carcinomas were found to be associated with CIMP-positive phenotype, which has been associated with worse prognosis of colorectal cancer irrespective of MSI status
- The profile of somatic mutations in LLS colorectal carcinomas differed from that in LS colorectal carcinomas
- Frequent *PIK3CA* mutations may predict resistance to anti-EGFR therapy

Study III

- Approximately 50% of LS breast carcinomas were MMR-deficient, as a marked distinction to sporadic breast carcinomas

- MMR-deficient LS breast carcinomas showed mainly hypermutated phenotype (91%), characteristic for other LS tumors, whereas MMR-proficient LS breast carcinomas and breast carcinomas from non-carrier family members were mainly non-hypermutated
- Mutational signatures of MMR-deficient breast carcinomas resembled the signatures of other LS tumors (LS colorectal and ovarian) and differed from breast carcinomas from non-carrier family members, but MMR-proficient LS breast carcinomas, too, exhibited mutational signatures associated with MMR-deficiency

Study IV

- Genetic and epigenetic characteristics of synchronous LS ovarian and endometrial tumors indicated shared origin, that is, they are very likely a primary tumor and its metastatic tumor
- Long term retrospective surveillance of female LS patients with endometrial hyperplasias and endometrial and ovarian carcinomas as endpoint lesions showed a high degree of genetic and epigenetic similarity between the specimens. This suggested early convergence of ovarian and endometrial tumorigenesis, which should be taken into account whenever ovarian or endometrial lesions are observed

ACKNOWLEDGEMENTS

REFERENCES

Cancer Gene Census. https://cancer.sanger.ac.uk/census. (Accessed Aug 20 2019).

Cancer Society of Finland, All about cancer. https://www.allaboutcancer.fi/. (Accessed Nov 22 2019).

Catalogue of somatic mutations in cancer (COSMIC). http://grch37-cancer.sanger.ac.uk/cosmic. (Accessed June 27 2019).

Database of Single Nucleotide Polymorphisms (dbSNP). https://www.ncbi.nlm.nih.gov/snp. (Accessed March 28 2017).

FATHMM. http://fathmm.biocompute.org.uk/. (Accessed August 25 2017).

Finnish Cancer registry. https://cancerregistry.fi/statistics/cancer-statistics/. (Accessed 27 Nov 2019).

GeneCards, Human Gene Database. http://www.genecards.org. (Accessed Sep 24 2019).

Genome Aggregation Database, gnomAD. https://gnomad.broadinstitute.org/. (Accessed Dec 13 2019).

GLOBOCAN 2012 v1.1, Cancer Incidence and Mortality Worldwide. (Accessed September 2018).

Human Splicing Finder. http://www.umd.be/HSF3/. (Accessed June 5 2017).

InSIGHT database (Leiden Open Variation Database, LOVD v. 2.0 Build 36. http://chromium.lovd.nl/LOVD2/colon_cancer/home.php. (Accessed August 25 2017).

InSIGHT DNA Variant Database. http://insight-database.org/. (Accessed 28 Nov 2019).

International Agency for Research on Cancer. http://gco.iarc.fr/today/home. (Accessed 26 Nov 2019).

MRC Holland. http://mrc-holland.com. (Accessed Aug 20 2019).

MutationAssessor. http://mutationassessor.org/r3/. (Accessed August 25 2017).

MuTect2. https://software.broadinstitute.org/gatk/documentation/tooldocs/3.8-0/org_broadinstitute_gatk_tools_walkers_cancer_m2_MuTect2.php. (Accessed Oct 2019 2019).

NCBI Gene tests database. https://www.ncbi.nlm.nih.gov/. (Accessed Sep 20 2019).

PubMed. https://www.ncbi.nlm.nih.gov/pubmed. (Accessed Feb 06 2020).

Sanger institute Cancer Gene Census. https://www.sanger.ac.uk/science/data/cancer-gene-census. (Accessed Feb 06 2020).

SIFT. http://sift.jcvi.org/. (Accessed August 25 2017).

Varsome - The Human Genomics Community. https://varsome.com/. (Accessed Nov 11 2019).

Aaltonen, L.A., Salovaara, R., Kristo, P., Canzian, F., Hemminki, A., Peltomaki, P., Chadwick, R.B., Kaariainen, H., Eskelinen, M., Jarvinen, H., Mecklin, J.P., de la Chapelle, A., 1998. Incidence of hereditary nonpolyposis colorectal cancer and the feasibility of molecular screening for the disease. N Engl J Med 338(21), 1481-1487.

Aarnio, M., Sankila, R., Pukkala, E., Salovaara, R., Aaltonen, L.A., de la Chapelle, A., Peltomaki, P., Mecklin, J.P., Jarvinen, H.J., 1999. Cancer risk in mutation carriers of DNA-mismatch-repair genes. Int J Cancer 81(2), 214-218.

Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., Sunyaev, S.R., 2010. A method and server for predicting damaging missense mutations. Nat Methods 7(4), 248-249.

AlDubayan, S.H., Giannakis, M., Moore, N.D., Han, G.C., Reardon, B., Hamada, T., Mu, X.J., Nishihara, R., Qian, Z., Liu, L., Yurgelun, M.B., Syngal, S., Garraway, L.A., Ogino, S., Fuchs, C.S., Van Allen, E.M., 2018. Inherited DNA-Repair Defects in Colorectal Cancer. Am J Hum Genet 102(3), 401-414.

Andor, N., Graham, T.A., Jansen, M., Xia, L.C., Aktipis, C.A., Petritsch, C., Ji, H.P., Maley, C.C., 2016. Pan-cancer analysis of the extent and consequences of intratumor heterogeneity. Nat Med 22(1), 105-113.

Anglesio, M.S., Wang, Y.K., Maassen, M., Horlings, H.M., Bashashati, A., Senz, J., Mackenzie, R., Grewal, D.S., Li-Chang, H., Karnezis, A.N., Sheffield, B.S., McConechy, M.K., Kommoss, F., Taran, F.A., Staebler, A., Shah, S.P., Wallwiener, D., Brucker, S., Gilks, C.B., Kommoss, S., Huntsman, D.G., 2016. Synchronous Endometrial and Ovarian Carcinomas: Evidence of Clonality. J Natl Cancer Inst 108(6), djv428.

Armaghany, T., Wilson, J.D., Chu, Q., Mills, G., 2012. Genetic alterations in colorectal cancer. Gastrointest Cancer Res 5(1), 19-27.

Atchley, D.P., Albarracin, C.T., Lopez, A., Valero, V., Amos, C.I., Gonzalez-Angulo, A.M., Hortobagyi, G.N., Arun, B.K., 2008. Clinical and pathologic characteristics of patients with BRCA-positive and BRCA-negative breast cancer. J Clin Oncol 26(26), 4282-4288.

Baylin, S.B., Hoppener, J.W., de Bustros, A., Steenbergh, P.H., Lips, C.J., Nelkin, B.D., 1986. DNA methylation patterns of the calcitonin gene in human lung cancers and lymphomas. Cancer Res 46(6), 2917-2922.

Berg, M., Hagland, H.R., Soreide, K., 2014. Comparison of CpG island methylator phenotype (CIMP) frequency in colon cancer using different probe- and gene-specific scoring alternatives on recommended multi-gene panels. PLoS One 9(1), e86657.

Beroud, C., Letovsky, S.I., Braastad, C.D., Caputo, S.M., Beaudoux, O., Bignon, Y.J., Bressac-De Paillerets, B., Bronner, M., Buell, C.M., Collod-Beroud, G., Coulet, F., Derive, N., Divincenzo, C., Elzinga, C.D., Garrec, C., Houdayer, C., Karbassi, I., Lizard, S., Love, A.,

Muller, D., Nagan, N., Nery, C.R., Rai, G., Revillion, F., Salgado, D., Sevenet, N., Sinilnikova, O., Sobol, H., Stoppa-Lyonnet, D., Toulas, C., Trautman, E., Vaur, D., Vilquin, P., Weymouth, K.S., Willis, A., Laboratory Corporation of America Variant Classification, G., Quest Diagnostics Variant Classification, G., Network, U.G.G.B.L., Eisenberg, M., Strom, C.M., 2016. BRCA Share: A Collection of Clinical BRCA Gene Variants. Hum Mutat 37(12), 1318-1328.

Bettstetter, M., Dechant, S., Ruemmele, P., Vogel, C., Kurz, K., Morak, M., Keller, G., Holinski-Feder, E., Hofstaedter, F., Dietmaier, W., 2008. MethyQESD, a robust and fast method for quantitative methylation analyses in HNPCC diagnostics using formalin-fixed and paraffin-embedded tissue samples. Lab Invest 88(12), 1367-1375.

Blokzijl, F., Janssen, R., van Boxtel, R., Cuppen, E., 2018. MutationalPatterns: comprehensive genome-wide analysis of mutational processes. Genome Med 10(1), 33.

Boland, C.R., Goel, A., 2010. Microsatellite instability in colorectal cancer. Gastroenterology 138(6), 2073-2087 e2073.

Boland, C.R., Thibodeau, S.N., Hamilton, S.R., Sidransky, D., Eshleman, J.R., Burt, R.W., Meltzer, S.J., Rodriguez-Bigas, M.A., Fodde, R., Ranzani, G.N., Srivastava, S., 1998. A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. Cancer Res 58(22), 5248-5257.

Bonadona, V., Bonaiti, B., Olschwang, S., Grandjouan, S., Huiart, L., Longy, M., Guimbaud, R., Buecher, B., Bignon, Y.J., Caron, O., Colas, C., Nogues, C., Lejeune-Dumoulin, S., Olivier-Faivre, L., Polycarpe-Osaer, F., Nguyen, T.D., Desseigne, F., Saurin, J.C., Berthet, P., Leroux, D., Duffour, J., Manouvrier, S., Frebourg, T., Sobol, H., Lasset, C., Bonaiti-Pellie, C., French Cancer Genetics, N., 2011. Cancer risks associated with germline mutations in MLH1, MSH2, and MSH6 genes in Lynch syndrome. JAMA 305(22), 2304-2310.

Buerki, N., Gautier, L., Kovac, M., Marra, G., Buser, M., Mueller, H., Heinimann, K., 2012. Evidence for breast cancer as an integral part of Lynch syndrome. Genes Chromosomes Cancer 51(1), 83-91.

Cancer Genome Atlas, N., 2012. Comprehensive molecular characterization of human colon and rectal cancer. Nature 487(7407), 330-337.

Carethers, J.M., 2014. Differentiating Lynch-like from Lynch syndrome. Gastroenterology 146(3), 602-604.

Choukrallah, M.A., Martin, F., Sierro, N., Hoeng, J., Ivanov, N.V., Peitsch, M.C., 2019. Computational Epigenomics: From Fundamental Research to Disease Prediction and Risk Assessment, in: Husi, H. (Ed.) Computational Biology. Brisbane (AU).

Church, D.N., Stelloo, E., Nout, R.A., Valtcheva, N., Depreeuw, J., ter Haar, N., Noske, A., Amant, F., Tomlinson, I.P., Wild, P.J., Lambrechts, D., Jurgenliemk-Schulz, I.M., Jobsen, J.J., Smit, V.T., Creutzberg, C.L., Bosse, T., 2015. Prognostic significance of POLE proofreading mutations in endometrial cancer. J Natl Cancer Inst 107(1), 402.

Cingolani, P., Platts, A., Wang le, L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., Ruden, D.M., 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) 6(2), 80-92.

Cohen, S.A., Turner, E.H., Beightol, M.B., Jacobson, A., Gooley, T.A., Salipante, S.J., Haraldsdottir, S., Smith, C., Scroggins, S., Tait, J.F., Grady, W.M., Lin, E.H., Cohn, D.E., Goodfellow, P.J., Arnold, M.W., de la Chapelle, A., Pearlman, R., Hampel, H., Pritchard, C.C., 2016. Frequent PIK3CA Mutations in Colorectal and Endometrial Tumors With 2 or More Somatic Mutations in Mismatch Repair Genes. Gastroenterology 151(3), 440-447 e441.

Couch, F.J., Nathanson, K.L., Offit, K., 2014. Two decades after BRCA: setting paradigms in personalized cancer care and prevention. Science 343(6178), 1466-1470.

Croce, C.M., 2008. Oncogenes and cancer. N Engl J Med 358(5), 502-511.

Cross, S.H., Bird, A.P., 1995. CpG islands and genes. Curr Opin Genet Dev 5(3), 309-314.

Cutting, G.R., 2014. Annotating DNA variants is the next major goal for human genetics. Am J Hum Genet 94(1), 5-10.

Davies, H., Morganella, S., Purdie, C.A., Jang, S.J., Borgen, E., Russnes, H., Glodzik, D., Zou, X., Viari, A., Richardson, A.L., Borresen-Dale, A.L., Thompson, A., Eyfjord, J.E., Kong, G., Stratton, M.R., Nik-Zainal, S., 2017. Whole-Genome Sequencing Reveals Breast Cancers with Mismatch Repair Deficiency. Cancer Res 77(18), 4755-4762.

Dermadi, D., Valo, S., Pussila, M., Reyhani, N., Sarantaus, L., Lalowski, M., Baumann, M., Nystrom, M., 2014. Inherited cancer predisposition sensitizes colonic mucosa to address Western diet effects and putative cancer-predisposing changes on mouse proteome. J Nutr Biochem 25(11), 1196-1206.

DeRycke, M.S., Gunawardena, S.R., Middha, S., Asmann, Y.W., Schaid, D.J., McDonnell, S.K., Riska, S.M., Eckloff, B.W., Cunningham, J.M., Fridley, B.L., Serie, D.J., Bamlet, W.R., Cicek, M.S., Jenkins, M.A., Duggan, D.J., Buchanan, D., Clendenning, M., Haile, R.W., Woods, M.O., Gallinger, S.N., Casey, G., Potter, J.D., Newcomb, P.A., Le Marchand, L., Lindor, N.M., Thibodeau, S.N., Goode, E.L., 2013. Identification of novel variants in colorectal cancer families by high-throughput exome sequencing. Cancer Epidemiol Biomarkers Prev 22(7), 1239-1251.

Desmond, A., Kurian, A.W., Gabree, M., Mills, M.A., Anderson, M.J., Kobayashi, Y., Horick, N., Yang, S., Shannon, K.M., Tung, N., Ford, J.M., Lincoln, S.E., Ellisen, L.W., 2015. Clinical Actionability of Multigene Panel Testing for Hereditary Breast and Ovarian Cancer Risk Assessment. JAMA Oncol 1(7), 943-951.

Dominguez-Valentin, M., Sampson, J.R., Seppala, T.T., Ten Broeke, S.W., Plazzer, J.P., Nakken, S., Engel, C., Aretz, S., Jenkins, M.A., Sunde, L., Bernstein, I., Capella, G., Balaguer, F., Thomas, H., Evans, D.G., Burn, J., Greenblatt, M., Hovig, E., de Vos Tot Nederveen Cappel, W.H., Sijmons, R.H., Bertario, L., Tibiletti, M.G., Cavestro, G.M., Lindblom, A., Della Valle, A., Lopez-Kostner, F., Gluck, N., Katz, L.H., Heinimann, K., Vaccaro, C.A., Buttner, R., Gorgens, H., Holinski-Feder, E., Morak, M., Holzapfel, S., Huneburg, R., Knebel Doeberitz, M.V., Loeffler, M., Rahner, N., Schackert, H.K., Steinke-Lange, V., Schmiegel, W., Vangala,

D., Pylvanainen, K., Renkonen-Sinisalo, L., Hopper, J.L., Win, A.K., Haile, R.W., Lindor, N.M., Gallinger, S., Le Marchand, L., Newcomb, P.A., Figueiredo, J.C., Thibodeau, S.N., Wadt, K., Therkildsen, C., Okkels, H., Ketabi, Z., Moreira, L., Sanchez, A., Serra-Burriel, M., Pineda, M., Navarro, M., Blanco, I., Green, K., Lalloo, F., Crosbie, E.J., Hill, J., Denton, O.G., Frayling, I.M., Rodland, E.A., Vasen, H., Mints, M., Neffa, F., Esperon, P., Alvarez, K., Kariv, R., Rosner, G., Pinero, T.A., Gonzalez, M.L., Kalfayan, P., Tjandra, D., Winship, I.M., Macrae, F., Moslein, G., Mecklin, J.P., Nielsen, M., Moller, P., 2019. Cancer risks by gene, age, and gender in 6350 carriers of pathogenic mismatch repair variants: findings from the Prospective Lynch Syndrome Database. Genet Med.

Elsayed, F.A., Kets, C.M., Ruano, D., van den Akker, B., Mensenkamp, A.R., Schrumpf, M., Nielsen, M., Wijnen, J.T., Tops, C.M., Ligtenberg, M.J., Vasen, H.F., Hes, F.J., Morreau, H., van Wezel, T., 2015. Germline variants in POLE are associated with early onset mismatch repair deficient colorectal cancer. Eur J Hum Genet 23(8), 1080-1084.

Engel, C., Loeffler, M., Steinke, V., Rahner, N., Holinski-Feder, E., Dietmaier, W., Schackert, H.K., Goergens, H., von Knebel Doeberitz, M., Goecke, T.O., Schmiegel, W., Buettner, R., Moeslein, G., Letteboer, T.G., Gomez Garcia, E., Hes, F.J., Hoogerbrugge, N., Menko, F.H., van Os, T.A., Sijmons, R.H., Wagner, A., Kluijt, I., Propping, P., Vasen, H.F., 2012. Risks of less common cancers in proven mutation carriers with lynch syndrome. J Clin Oncol 30(35), 4409-4415.

Esemuede, I., Forslund, A., Khan, S.A., Qin, L.X., Gimbel, M.I., Nash, G.M., Zeng, Z., Rosenberg, S., Shia, J., Barany, F., Paty, P.B., 2010. Improved testing for microsatellite instability in colorectal cancer using a simplified 3-marker assay. Ann Surg Oncol 17(12), 3370-3378.

Espenschied, C.R., LaDuca, H., Li, S., McFarland, R., Gau, C.L., Hampel, H., 2017. Multigene Panel Testing Provides a New Perspective on Lynch Syndrome. J Clin Oncol 35(22), 2568-2575.

Esteller, M., Herman, J.G., 2002. Cancer as an epigenetic disease: DNA methylation and chromatin alterations in human tumours. J Pathol 196(1), 1-7.

Fearon, E.R., Vogelstein, B., 1990. A genetic model for colorectal tumorigenesis. Cell 61(5), 759-767.

Feinberg, A.P., Tycko, B., 2004. The history of cancer epigenetics. Nat Rev Cancer 4(2), 143-153.

Futreal, P.A., Liu, Q., Shattuck-Eidens, D., Cochran, C., Harshman, K., Tavtigian, S., Bennett, L.M., Haugen-Strano, A., Swensen, J., Miki, Y., et al., 1994. BRCA1 mutations in primary breast and ovarian carcinomas. Science 266(5182), 120-122.

Geurts-Giele, W.R., Leenen, C.H., Dubbink, H.J., Meijssen, I.C., Post, E., Sleddens, H.F., Kuipers, E.J., Goverde, A., van den Ouweland, A.M., van Lier, M.G., Steyerberg, E.W., van Leerdam, M.E., Wagner, A., Dinjens, W.N., 2014. Somatic aberrations of mismatch repair genes as a cause of microsatellite-unstable cancers. J Pathol 234(4), 548-559.

Green, E.D., Watson, J.D., Collins, F.S., 2015. Human Genome Project: Twenty-five years of big biology. Nature 526(7571), 29-31.

Hampel, H., Frankel, W., Panescu, J., Lockman, J., Sotamaa, K., Fix, D., Comeras, I., La Jeunesse, J., Nakagawa, H., Westman, J.A., Prior, T.W., Clendenning, M., Penzone, P., Lombardi, J., Dunn, P., Cohn, D.E., Copeland, L., Eaton, L., Fowler, J., Lewandowski, G., Vaccarello, L., Bell, J., Reid, G., de la Chapelle, A., 2006. Screening for Lynch syndrome (hereditary nonpolyposis colorectal cancer) among endometrial cancer patients. Cancer Res 66(15), 7810-7817.

Hampel, H., Frankel, W.L., Martin, E., Arnold, M., Khanduja, K., Kuebler, P., Clendenning, M., Sotamaa, K., Prior, T., Westman, J.A., Panescu, J., Fix, D., Lockman, J., LaJeunesse, J., Comeras, I., de la Chapelle, A., 2008. Feasibility of screening for Lynch syndrome among patients with colorectal cancer. J Clin Oncol 26(35), 5783-5788.

Hanahan, D., Weinberg, R.A., 2000. The hallmarks of cancer. Cell 100(1), 57-70.

Hanahan, D., Weinberg, R.A., 2011. Hallmarks of cancer: the next generation. Cell 144(5), 646-674.

Haraldsdottir, S., Hampel, H., Tomsic, J., Frankel, W.L., Pearlman, R., de la Chapelle, A., Pritchard, C.C., 2014. Colon and endometrial cancers with mismatch repair deficiency can arise from somatic, rather than germline, mutations. Gastroenterology 147(6), 1308-1316 e1301.

Hauke, J., Hahnen, E., Schneider, S., Reuss, A., Richters, L., Kommoss, S., Heimbach, A., Marme, F., Schmidt, S., Prieske, K., Gevensleben, H., Burges, A., Borde, J., De Gregorio, N., Nurnberg, P., El-Balat, A., Thiele, H., Hilpert, F., Altmuller, J., Meier, W., Dietrich, D., Kimmig, R., Schoemig-Markiefka, B., Kast, K., Braicu, E., Baumann, K., Jackisch, C., Park-Simon, T.W., Ernst, C., Hanker, L., Pfisterer, J., Schnelzer, A., du Bois, A., Schmutzler, R.K., Harter, P., 2019. Deleterious somatic variants in 473 consecutive individuals with ovarian cancer: results of the observational AGO-TR1 study (NCT02222883). J Med Genet 56(9), 574-580.

Helder-Woolderink, J.M., Blok, E.A., Vasen, H.F., Hollema, H., Mourits, M.J., De Bock, G.H., 2016. Ovarian cancer in Lynch syndrome; a systematic review. Eur J Cancer 55, 65-73.

Hulka, B.S., 1997. Epidemiologic analysis of breast and gynecologic cancers. Prog Clin Biol Res 396, 17-29.

Imai, K., Yamamoto, H., 2008. Carcinogenesis and microsatellite instability: the interrelationship between genetics and epigenetics. Carcinogenesis 29(4), 673-680.

International Human Genome Sequencing, C., 2004. Finishing the euchromatic sequence of the human genome. Nature 431(7011), 931-945.

Isola, J., DeVries, S., Chu, L., Ghazvini, S., Waldman, F., 1994. Analysis of changes in DNA sequence copy number by comparative genomic hybridization in archival paraffin-embedded tumor samples. Am J Pathol 145(6), 1301-1308.

Jacob, S., Praz, F., 2002. DNA mismatch repair defects: role in colorectal carcinogenesis. Biochimie 84(1), 27-47.

Jansen, A.M., van Wezel, T., van den Akker, B.E., Ventayol Garcia, M., Ruano, D., Tops, C.M., Wagner, A., Letteboer, T.G., Gomez-Garcia, E.B., Devilee, P., Wijnen, J.T., Hes, F.J., Morreau, H., 2016. Combined mismatch repair and POLE/POLD1 defects explain unresolved suspected Lynch syndrome cancers. Eur J Hum Genet 24(7), 1089-1092.

Jeffreys, A.J., Wilson, V., Thein, S.L., 1992. Hypervariable 'minisatellite' regions in human DNA. 1985. Biotechnology 24, 467-472.

Jeggo, P.A., Pearl, L.H., Carr, A.M., 2016. DNA repair, genome stability and cancer: a historical perspective. Nat Rev Cancer 16(1), 35-42.

Jiang, X., Finucane, H.K., Schumacher, F.R., Schmit, S.L., Tyrer, J.P., Han, Y., Michailidou, K., Lesseur, C., Kuchenbaecker, K.B., Dennis, J., Conti, D.V., Casey, G., Gaudet, M.M., Huyghe, J.R., Albanes, D., Aldrich, M.C., Andrew, A.S., Andrulis, I.L., Anton-Culver, H., Antoniou, A.C., Antonenkova, N.N., Arnold, S.M., Aronson, K.J., Arun, B.K., Bandera, E.V., Barkardottir, R.B., Barnes, D.R., Batra, J., Beckmann, M.W., Benitez, J., Benlloch, S., Berchuck, A., Berndt, S.I., Bickeboller, H., Bien, S.A., Blomqvist, C., Boccia, S., Bogdanova, N.V., Bojesen, S.E., Bolla, M.K., Brauch, H., Brenner, H., Brenton, J.D., Brook, M.N., Brunet, J., Brunnstrom, H., Buchanan, D.D., Burwinkel, B., Butzow, R., Cadoni, G., Caldes, T., Caligo, M.A., Campbell, I., Campbell, P.T., Cancel-Tassin, G., Cannon-Albright, L., Campa, D., Caporaso, N., Carvalho, A.L., Chan, A.T., Chang-Claude, J., Chanock, S.J., Chen, C., Christiani, D.C., Claes, K.B.M., Claessens, F., Clements, J., Collee, J.M., Correa, M.C., Couch, F.J., Cox, A., Cunningham, J.M., Cybulski, C., Czene, K., Daly, M.B., deFazio, A., Devilee, P., Diez, O., Gago-Dominguez, M., Donovan, J.L., Dork, T., Duell, E.J., Dunning, A.M., Dwek, M., Eccles, D.M., Edlund, C.K., Edwards, D.R.V., Ellberg, C., Evans, D.G., Fasching, P.A., Ferris, R.L., Liloglou, T., Figueiredo, J.C., Fletcher, O., Fortner, R.T., Fostira, F., Franceschi, S., Friedman, E., Gallinger, S.J., Ganz, P.A., Garber, J., Garcia-Saenz, J.A., Gayther, S.A., Giles, G.G., Godwin, A.K., Goldberg, M.S., Goldgar, D.E., Goode, E.L., Goodman, M.T., Goodman, G., Grankvist, K., Greene, M.H., Gronberg, H., Gronwald, J., Guenel, P., Hakansson, N., Hall, P., Hamann, U., Hamdy, F.C., Hamilton, R.J., Hampe, J., Haugen, A., Heitz, F., Herrero, R., Hillemanns, P., Hoffmeister, M., Hogdall, E., Hong, Y.C., Hopper, J.L., Houlston, R., Hulick, P.J., Hunter, D.J., Huntsman, D.G., Idos, G., Imyanitov, E.N., Ingles, S.A., Isaacs, C., Jakubowska, A., James, P., Jenkins, M.A., Johansson, M., Johansson, M., John, E.M., Joshi, A.D., Kaneva, R., Karlan, B.Y., Kelemen, L.E., Kuhl, T., Khaw, K.T., Khusnutdinova, E., Kibel, A.S., Kiemeney, L.A., Kim, J., Kjaer, S.K., Knight, J.A., Kogevinas, M., Kote-Jarai, Z., Koutros, S., Kristensen, V.N., Kupryjanczyk, J., Lacko, M., Lam, S., Lambrechts, D., Landi, M.T., Lazarus, P., Le, N.D., Lee, E., Lejbkowicz, F., Lenz, H.J., Leslie, G., Lessel, D., Lester, J., Levine, D.A., Li, L., Li, C.I., Lindblom, A., Lindor, N.M., Liu, G., Loupakis, F., Lubinski, J., Maehle, L., Maier, C., Mannermaa, A., Marchand, L.L., Margolin, S., May, T., McGuffog, L., Meindl, A., Middha, P., Miller, A., Milne, R.L., MacInnis, R.J., Modugno, F., Montagna, M., Moreno, V., Moysich, K.B., Mucci, L., Muir, K., Mulligan, A.M., Nathanson, K.L., Neal, D.E., Ness, A.R., Neuhausen, S.L., Nevanlinna, H., Newcomb, P.A., Newcomb, L.F., Nielsen, F.C., Nikitina-Zake, L., Nordestgaard, B.G., Nussbaum, R.L., Offit, K., Olah, E., Olama, A.A.A., Olopade, O.I., Olshan, A.F., Olsson, H., Osorio, A., Pandha, H., Park, J.Y., Pashayan, N., Parsons, M.T., Pejovic, T., Penney, K.L., Peters, W.H.M., Phelan, C.M., Phipps, A.I., Plaseska-Karanfilska, D., Pring, M., Prokofyeva, D., Radice, P., Stefansson, K., Ramus, S.J., Raskin, L., Rennert, G., Rennert, H.S., van Rensburg, E.J., Riggan, M.J., Risch, H.A., Risch, A., Roobol, M.J., Rosenstein, B.S., Rossing, M.A., De Ruyck, K., Saloustros, E., Sandler, D.P., Sawyer, E.J., Schabath, M.B., Schleutker, J., Schmidt, M.K., Setiawan, V.W., Shen, H., Siegel, E.M., Sieh, W., Singer, C.F., Slattery, M.L., Sorensen, K.D., Southey, M.C., Spurdle, A.B., Stanford, J.L., Stevens, V.L., Stintzing, S., Stone, J., Sundfeldt,

K., Sutphen, R., Swerdlow, A.J., Tajara, E.H., Tangen, C.M., Tardon, A., Taylor, J.A., Teare, M.D., Teixeira, M.R., Terry, M.B., Terry, K.L., Thibodeau, S.N., Thomassen, M., Bjorge, L., Tischkowitz, M., Toland, A.E., Torres, D., Townsend, P.A., Travis, R.C., Tung, N., Tworoger, S.S., Ulrich, C.M., Usmani, N., Vachon, C.M., Van Nieuwenhuysen, E., Vega, A., Aguado-Barrera, M.E., Wang, Q., Webb, P.M., Weinberg, C.R., Weinstein, S., Weissler, M.C., Weitzel, J.N., West, C.M.L., White, E., Whittemore, A.S., Wichmann, H.E., Wiklund, F., Winqvist, R., Wolk, A., Woll, P., Woods, M., Wu, A.H., Wu, X., Yannoukakos, D., Zheng, W., Zienolddiny, S., Ziogas, A., Zorn, K.K., Lane, J.M., Saxena, R., Thomas, D., Hung, R.J., Diergaarde, B., McKay, J., Peters, U., Hsu, L., Garcia-Closas, M., Eeles, R.A., Chenevix-Trench, G., Brennan, P.J., Haiman, C.A., Simard, J., Easton, D.F., Gruber, S.B., Pharoah, P.D.P., Price, A.L., Pasaniuc, B., Amos, C.I., Kraft, P., Lindstrom, S., 2019. Shared heritability and functional enrichment across six solid cancers. Nat Commun 10(1), 431.

Joensuu, E.I., Abdel-Rahman, W.M., Ollikainen, M., Ruosaari, S., Knuutila, S., Peltomaki, P., 2008. Epigenetic signatures of familial cancer are characteristic of tumor type and family category. Cancer Res 68(12), 4597-4605.

Jones, P.A., 2012. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet 13(7), 484-492.

Juvet, L.K., Natvig Norderhaug, I., 2008. Genetic Tests for Breast and Ovarian Cancer. Oslo, Norway.

Kansikas, M., Kariola, R., Nystrom, M., 2011. Verification of the three-step model in assessing the pathogenicity of mismatch repair gene variants. Hum Mutat 32(1), 107-115.

Karnezis, A.N., Cho, K.R., Gilks, C.B., Pearce, C.L., Huntsman, D.G., 2017. The disparate origins of ovarian cancers: pathogenesis and prevention strategies. Nat Rev Cancer 17(1), 65-74.

Kasela, M., Nystrom, M., Kansikas, M., 2019. PMS2 expression decrease causes severe problems in mismatch repair. Hum Mutat 40(7), 904-907.

Kelemen, L.E., Rambau, P.F., Koziak, J.M., Steed, H., Kobel, M., 2017. Synchronous endometrial and ovarian carcinomas: predictors of risk and associations with survival and tumor expression profiles. Cancer Causes Control 28(5), 447-457.

Knudson, A.G., Jr., 1971. Mutation and cancer: statistical study of retinoblastoma. Proc Natl Acad Sci U S A 68(4), 820-823.

Koboldt, D.C., Zhang, Q., Larson, D.E., Shen, D., McLellan, M.D., Lin, L., Miller, C.A., Mardis, E.R., Ding, L., Wilson, R.K., 2012. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res 22(3), 568-576.

Kok, M., Horlings, H.M., Snaebjornsson, P., Chalabi, M., Schumacher, T.N., Blank, C.U., Linn, S.C., van Dieren, J., 2017. Case report: Profound Immunotherapy Response in Mismatch Repair-Deficient Breast Cancer, JCO™ Precision Oncology. American Society of Clinical Oncology, ascopubs.org/journal/po, pp. 1-3.

Kuiper, R.P., Vissers, L.E., Venkatachalam, R., Bodmer, D., Hoenselaar, E., Goossens, M., Haufe, A., Kamping, E., Niessen, R.C., Hogervorst, F.B., Gille, J.J., Redeker, B., Tops, C.M.,

van Gijn, M.E., van den Ouweland, A.M., Rahner, N., Steinke, V., Kahl, P., Holinski-Feder, E., Morak, M., Kloor, M., Stemmler, S., Betz, B., Hutter, P., Bunyan, D.J., Syngal, S., Culver, J.O., Graham, T., Chan, T.L., Nagtegaal, I.D., van Krieken, J.H., Schackert, H.K., Hoogerbrugge, N., van Kessel, A.G., Ligtenberg, M.J., 2011. Recurrence and variability of germline EPCAM deletions in Lynch syndrome. Hum Mutat 32(4), 407-414.

Kurman, R.J., Shih Ie, M., 2008. Pathogenesis of ovarian cancer: lessons from morphology and molecular biology and their clinical implications. Int J Gynecol Pathol 27(2), 151-160.

Lahiri, D.K., Nurnberger, J.I., Jr., 1991. A rapid non-enzymatic method for the preparation of HMW DNA from blood for RFLP studies. Nucleic Acids Res 19(19), 5444.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J.P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, Y., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J.C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R.H., Wilson, R.K., Hillier, L.W., McPherson, J.D., Marra, M.A., Mardis, E.R., Fulton, L.A., Chinwalla, A.T., Pepin, K.H., Gish, W.R., Chissoe, S.L., Wendl, M.C., Delehaunty, K.D., Miner, T.L., Delehaunty, A., Kramer, J.B., Cook, L.L., Fulton, R.S., Johnson, D.L., Minx, P.J., Clifton, S.W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J.F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Gibbs, R.A., Muzny, D.M., Scherer, S.E., Bouck, J.B., Sodergren, E.J., Worley, K.C., Rives, C.M., Gorrell, J.H., Metzker, M.L., Naylor, S.L., Kucherlapati, R.S., Nelson, D.L., Weinstock, G.M., Sakaki, Y., Fujiyama, A., Hattori, M., Yada, T., Toyoda, A., Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., Weissenbach, J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Bruls, T., Pelletier, E., Robert, C., Wincker, P., Smith, D.R., Doucette-Stamm, L., Rubenfield, M., Weinstock, K., Lee, H.M., Dubois, J., Rosenthal, A., Platzer, M., Nyakatura, G., Taudien, S., Rump, A., Yang, H., Yu, J., Wang, J., Huang, G., Gu, J., Hood, L., Rowen, L., Madan, A., Qin, S., Davis, R.W., Federspiel, N.A., Abola, A.P., Proctor, M.J., Myers, R.M., Schmutz, J., Dickson, M., Grimwood, J., Cox, D.R., Olson, M.V., Kaul, R., Raymond, C., Shimizu, N., Kawasaki, K., Minoshima, S., Evans, G.A., Athanasiou, M., Schultz, R., Roe, B.A., Chen, F., Pan, H., Ramser, J., Lehrach, H., Reinhardt, R., McCombie, W.R., de la Bastide, M., Dedhia, N., Blocker, H., Hornischer, K., Nordsiek, G., Agarwala, R., Aravind, L., Bailey, J.A., Bateman, A., Batzoglou, S., Birney, E., Bork, P., Brown, D.G., Burge, C.B., Cerutti, L., Chen, H.C., Church, D., Clamp, M., Copley, R.R., Doerks, T., Eddy, S.R., Eichler, E.E., Furey, T.S., Galagan, J., Gilbert, J.G., Harmon, C., Hayashizaki, Y., Haussler, D., Hermjakob, H., Hokamp, K., Jang, W., Johnson, L.S., Jones, T.A., Kasif, S., Kaspryzk, A., Kennedy, S., Kent, W.J., Kitts, P., Koonin, E.V., Korf, I., Kulp, D., Lancet, D., Lowe, T.M., McLysaght, A., Mikkelsen, T., Moran, J.V., Mulder, N., Pollara, V.J., Ponting, C.P., Schuler, G., Schultz, J., Slater, G., Smit, A.F., Stupka, E., Szustakowki, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., Wheeler, R., Williams, A., Wolf, Y.I., Wolfe, K.H., Yang, S.P., Yeh, R.F., Collins, F., Guyer, M.S., Peterson, J., Felsenfeld, A., Wetterstrand, K.A., Patrinos, A., Morgan, M.J., de Jong, P., Catanese, J.J., Osoegawa, K., Shizuya, H., Choi, S., Chen, Y.J., Szustakowki, J., International Human Genome Sequencing, C., 2001. Initial sequencing and analysis of the human genome. Nature 409(6822), 860-921.

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25(14), 1754-1760.

Li, Y.C., Korol, A.B., Fahima, T., Beiles, A., Nevo, E., 2002. Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. Mol Ecol 11(12), 2453-2465.

Liu, D., Keijzers, G., Rasmussen, L.J., 2017. DNA mismatch repair and its many roles in eukaryotic cells. Mutat Res 773, 174-187.

Loeb, L.A., 1991. Mutator phenotype may be required for multistage carcinogenesis. Cancer Res 51(12), 3075-3079.

Lotsari, J.E., Gylling, A., Abdel-Rahman, W.M., Nieminen, T.T., Aittomaki, K., Friman, M., Pitkanen, R., Aarnio, M., Jarvinen, H.J., Mecklin, J.P., Kuopio, T., Peltomaki, P., 2012. Breast carcinoma and Lynch syndrome: molecular analysis of tumors arising in mutation carriers, non-carriers, and sporadic cases. Breast Cancer Res 14(3), R90.

Loukola, A., Eklin, K., Laiho, P., Salovaara, R., Kristo, P., Jarvinen, H., Mecklin, J.P., Launonen, V., Aaltonen, L.A., 2001. Microsatellite marker analysis in screening for hereditary nonpolyposis colorectal cancer (HNPCC). Cancer Res 61(11), 4545-4549.

Lynch, H.T., de la Chapelle, A., 2003. Hereditary colorectal cancer. N Engl J Med 348(10), 919-932.

Lynch, H.T., Lynch, J.F., Lynch, P.M., 2007. Toward a consensus in molecular diagnosis of hereditary nonpolyposis colorectal cancer (Lynch syndrome). J Natl Cancer Inst 99(4), 261-263.

Malander, S., Rambech, E., Kristoffersson, U., Halvarsson, B., Ridderheim, M., Borg, A., Nilbert, M., 2006. The contribution of the hereditary nonpolyposis colorectal cancer syndrome to the development of ovarian cancer. Gynecol Oncol 101(2), 238-243.

Mas-Moya, J., Dudley, B., Brand, R.E., Thull, D., Bahary, N., Nikiforova, M.N., Pai, R.K., 2015. Clinicopathological comparison of colorectal and endometrial carcinomas in patients with Lynch-like syndrome versus patients with Lynch syndrome. Hum Pathol 46(11), 1616-1625.

Mavaddat, N., Barrowdale, D., Andrulis, I.L., Domchek, S.M., Eccles, D., Nevanlinna, H., Ramus, S.J., Spurdle, A., Robson, M., Sherman, M., Mulligan, A.M., Couch, F.J., Engel, C., McGuffog, L., Healey, S., Sinilnikova, O.M., Southey, M.C., Terry, M.B., Goldgar, D., O'Malley, F., John, E.M., Janavicius, R., Tihomirova, L., Hansen, T.V., Nielsen, F.C., Osorio, A., Stavropoulou, A., Benitez, J., Manoukian, S., Peissel, B., Barile, M., Volorio, S., Pasini, B., Dolcetti, R., Putignano, A.L., Ottini, L., Radice, P., Hamann, U., Rashid, M.U., Hogervorst, F.B., Kriege, M., van der Luijt, R.B., Hebon, Peock, S., Frost, D., Evans, D.G., Brewer, C., Walker, L., Rogers, M.T., Side, L.E., Houghton, C., Embrace, Weaver, J., Godwin, A.K., Schmutzler, R.K., Wappenschmidt, B., Meindl, A., Kast, K., Arnold, N., Niederacher, D., Sutter, C., Deissler, H., Gadzicki, D., Preisler-Adams, S., Varon-Mateeva, R., Schonbuchner, I., Gevensleben, H., Stoppa-Lyonnet, D., Belotti, M., Barjhoux, L., Collaborators, G.S., Isaacs, C., Peshkin, B.N., Caldes, T., de la Hoya, M., Canadas, C., Heikkinen, T., Heikkila, P., Aittomaki, K., Blanco, I., Lazaro, C., Brunet, J., Agnarsson, B.A., Arason, A., Barkardottir, R.B., Dumont, M., Simard, J., Montagna, M., Agata, S., D'Andrea, E., Yan, M., Fox, S., kConFab, I., Rebbeck,

T.R., Rubinstein, W., Tung, N., Garber, J.E., Wang, X., Fredericksen, Z., Pankratz, V.S., Lindor, N.M., Szabo, C., Offit, K., Sakr, R., Gaudet, M.M., Singer, C.F., Tea, M.K., Rappaport, C., Mai, P.L., Greene, M.H., Sokolenko, A., Imyanitov, E., Toland, A.E., Senter, L., Sweet, K., Thomassen, M., Gerdes, A.M., Kruse, T., Caligo, M., Aretini, P., Rantala, J., von Wachenfeld, A., Henriksson, K., Collaborators, S.-B., Steele, L., Neuhausen, S.L., Nussbaum, R., Beattie, M., Odunsi, K., Sucheston, L., Gayther, S.A., Nathanson, K., Gross, J., Walsh, C., Karlan, B., Chenevix-Trench, G., Easton, D.F., Antoniou, A.C., Consortium of Investigators of Modifiers of, B., 2012. Pathology of breast and ovarian cancers among BRCA1 and BRCA2 mutation carriers: results from the Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA). Cancer Epidemiol Biomarkers Prev 21(1), 134-147.

McLaughlin, J.R., Rosen, B., Moody, J., Pal, T., Fan, I., Shaw, P.A., Risch, H.A., Sellers, T.A., Sun, P., Narod, S.A., 2013. Long-term ovarian cancer survival associated with mutation in BRCA1 or BRCA2. J Natl Cancer Inst 105(2), 141-148.

Mensenkamp, A.R., Vogelaar, I.P., van Zelst-Stams, W.A., Goossens, M., Ouchene, H., Hendriks-Cornelissen, S.J., Kwint, M.P., Hoogerbrugge, N., Nagtegaal, I.D., Ligtenberg, M.J., 2014. Somatic mutations in MLH1 and MSH2 are a frequent cause of mismatch-repair deficiency in Lynch syndrome-like tumors. Gastroenterology 146(3), 643-646 e648.

Minion, L.E., Dolinsky, J.S., Chase, D.M., Dunlop, C.L., Chao, E.C., Monk, B.J., 2015. Hereditary predisposition to ovarian cancer, looking beyond BRCA1/BRCA2. Gynecol Oncol 137(1), 86-92.

Moller, P., Seppala, T., Bernstein, I., Holinski-Feder, E., Sala, P., Evans, D.G., Lindblom, A., Macrae, F., Blanco, I., Sijmons, R., Jeffries, J., Vasen, H., Burn, J., Nakken, S., Hovig, E., Rodland, E.A., Tharmaratnam, K., de Vos Tot Nederveen Cappel, W.H., Hill, J., Wijnen, J., Green, K., Lalloo, F., Sunde, L., Mints, M., Bertario, L., Pineda, M., Navarro, M., Morak, M., Renkonen-Sinisalo, L., Frayling, I.M., Plazzer, J.P., Pylvanainen, K., Sampson, J.R., Capella, G., Mecklin, J.P., Moslein, G., Mallorca, G., 2017a. Cancer incidence and survival in Lynch syndrome patients receiving colonoscopic and gynaecological surveillance: first report from the prospective Lynch syndrome database. Gut 66(3), 464-472.

Moller, P., Seppala, T., Bernstein, I., Holinski-Feder, E., Sala, P., Evans, D.G., Lindblom, A., Macrae, F., Blanco, I., Sijmons, R., Jeffries, J., Vasen, H., Burn, J., Nakken, S., Hovig, E., Rodland, E.A., Tharmaratnam, K., de Vos Tot Nederveen Cappel, W.H., Hill, J., Wijnen, J., Jenkins, M., Green, K., Lalloo, F., Sunde, L., Mints, M., Bertario, L., Pineda, M., Navarro, M., Morak, M., Renkonen-Sinisalo, L., Frayling, I.M., Plazzer, J.P., Pylvanainen, K., Genuardi, M., Mecklin, J.P., Moslein, G., Sampson, J.R., Capella, G., Mallorca, G., 2017b. Incidence of and survival after subsequent cancers in carriers of pathogenic MMR variants with previous cancer: a report from the prospective Lynch syndrome database. Gut 66(9), 1657-1664.

Moller, P., Seppala, T.T., Bernstein, I., Holinski-Feder, E., Sala, P., Gareth Evans, D., Lindblom, A., Macrae, F., Blanco, I., Sijmons, R.H., Jeffries, J., Vasen, H.F.A., Burn, J., Nakken, S., Hovig, E., Rodland, E.A., Tharmaratnam, K., de Vos Tot Nederveen Cappel, W.H., Hill, J., Wijnen, J.T., Jenkins, M.A., Green, K., Lalloo, F., Sunde, L., Mints, M., Bertario, L., Pineda, M., Navarro, M., Morak, M., Renkonen-Sinisalo, L., Valentin, M.D., Frayling, I.M., Plazzer, J.P., Pylvanainen, K., Genuardi, M., Mecklin, J.P., Moeslein, G., Sampson, J.R., Capella, G., Mallorca, G., 2018. Cancer risk and survival in path_MMR carriers by gene and gender up to 75 years of age: a report from the Prospective Lynch Syndrome Database. Gut 67(7), 1306-1316.

Moreira, L., Balaguer, F., Lindor, N., de la Chapelle, A., Hampel, H., Aaltonen, L.A., Hopper, J.L., Le Marchand, L., Gallinger, S., Newcomb, P.A., Haile, R., Thibodeau, S.N., Gunawardena, S., Jenkins, M.A., Buchanan, D.D., Potter, J.D., Baron, J.A., Ahnen, D.J., Moreno, V., Andreu, M., Ponz de Leon, M., Rustgi, A.K., Castells, A., Consortium, E., 2012. Identification of Lynch syndrome among patients with colorectal cancer. JAMA 308(15), 1555-1565.

Morganti, S., Tarantino, P., Ferraro, E., D'Amico, P., Viale, G., Trapani, D., Duso, B.A., Curigliano, G., 2019. Complexity of genome sequencing and reporting: Next generation sequencing (NGS) technologies and implementation of precision medicine in real life. Crit Rev Oncol Hematol 133, 171-182.

Mork, M.E., You, Y.N., Ying, J., Bannon, S.A., Lynch, P.M., Rodriguez-Bigas, M.A., Vilar, E., 2015. High Prevalence of Hereditary Cancer Syndromes in Adolescents and Young Adults With Colorectal Cancer. J Clin Oncol 33(31), 3544-3549.

Mucci, L.A., Hjelmborg, J.B., Harris, J.R., Czene, K., Havelick, D.J., Scheike, T., Graff, R.E., Holst, K., Moller, S., Unger, R.H., McIntosh, C., Nuttall, E., Brandt, I., Penney, K.L., Hartman, M., Kraft, P., Parmigiani, G., Christensen, K., Koskenvuo, M., Holm, N.V., Heikkila, K., Pukkala, E., Skytthe, A., Adami, H.O., Kaprio, J., Nordic Twin Study of Cancer, C., 2016. Familial Risk and Heritability of Cancer Among Twins in Nordic Countries. JAMA 315(1), 68-76.

Niskakoski, A., Kaur, S., Renkonen-Sinisalo, L., Lassus, H., Jarvinen, H.J., Mecklin, J.P., Butzow, R., Peltomaki, P., 2013. Distinct molecular profiles in Lynch syndrome-associated and sporadic ovarian carcinomas. Int J Cancer 133(11), 2596-2608.

Norquist, B.M., Harrell, M.I., Brady, M.F., Walsh, T., Lee, M.K., Gulsuner, S., Bernards, S.S., Casadei, S., Yi, Q., Burger, R.A., Chan, J.K., Davidson, S.A., Mannel, R.S., DiSilvestro, P.A., Lankes, H.A., Ramirez, N.C., King, M.C., Swisher, E.M., Birrer, M.J., 2016. Inherited Mutations in Women With Ovarian Carcinoma. JAMA Oncol 2(4), 482-490.

Nowak, J.A., Yurgelun, M.B., Bruce, J.L., Rojas-Rudilla, V., Hall, D.L., Shivdasani, P., Garcia, E.P., Agoston, A.T., Srivastava, A., Ogino, S., Kuo, F.C., Lindeman, N.I., Dong, F., 2017. Detection of Mismatch Repair Deficiency and Microsatellite Instability in Colorectal Adenocarcinoma by Targeted Next-Generation Sequencing. J Mol Diagn 19(1), 84-91.

Nygren, A.O., Ameziane, N., Duarte, H.M., Vijzelaar, R.N., Waisfisz, Q., Hess, C.J., Schouten, J.P., Errami, A., 2005. Methylation-specific MLPA (MS-MLPA): simultaneous detection of CpG methylation and copy number changes of up to 40 sequences. Nucleic Acids Res 33(14), e128.

Nystrom-Lahti, M., Kristo, P., Nicolaides, N.C., Chang, S.Y., Aaltonen, L.A., Moisio, A.L., Jarvinen, H.J., Mecklin, J.P., Kinzler, K.W., Vogelstein, B., et al., 1995. Founding mutations and Alu-mediated recombination in hereditary colon cancer. Nat Med 1(11), 1203-1206.

Ogino, S., Kawasaki, T., Kirkner, G.J., Kraft, P., Loda, M., Fuchs, C.S., 2007. Evaluation of markers for CpG island methylator phenotype (CIMP) in colorectal cancer by a large population-based sample. J Mol Diagn 9(3), 305-314.

Ollikainen, M., Abdel-Rahman, W.M., Moisio, A.L., Lindroos, A., Kariola, R., Jarvela, I., Poyhonen, M., Butzow, R., Peltomaki, P., 2005. Molecular analysis of familial endometrial carcinoma: a manifestation of hereditary nonpolyposis colorectal cancer or a separate syndrome? J Clin Oncol 23(21), 4609-4616.

Palles, C., Cazier, J.B., Howarth, K.M., Domingo, E., Jones, A.M., Broderick, P., Kemp, Z., Spain, S.L., Guarino, E., Salguero, I., Sherborne, A., Chubb, D., Carvajal-Carmona, L.G., Ma, Y., Kaur, K., Dobbins, S., Barclay, E., Gorman, M., Martin, L., Kovac, M.B., Humphray, S., Consortium, C., Consortium, W.G.S., Lucassen, A., Holmes, C.C., Bentley, D., Donnelly, P., Taylor, J., Petridis, C., Roylance, R., Sawyer, E.J., Kerr, D.J., Clark, S., Grimes, J., Kearsey, S.E., Thomas, H.J., McVean, G., Houlston, R.S., Tomlinson, I., 2013. Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. Nat Genet 45(2), 136-144.

Palomaki, G.E., McClain, M.R., Melillo, S., Hampel, H.L., Thibodeau, S.N., 2009. EGAPP supplementary evidence review: DNA testing strategies aimed at reducing morbidity and mortality from Lynch syndrome. Genet Med 11(1), 42-65.

Pande, M., Wei, C., Chen, J., Amos, C.I., Lynch, P.M., Lu, K.H., Lucio, L.A., Boyd-Rogers, S.G., Bannon, S.A., Mork, M.E., Frazier, M.L., 2012. Cancer spectrum in DNA mismatch repair gene mutation carriers: results from a hospital based Lynch syndrome registry. Fam Cancer 11(3), 441-447.

Pavicic, W., Perkio, E., Kaur, S., Peltomaki, P., 2011. Altered methylation at microRNA-associated CpG islands in hereditary and sporadic carcinomas: a methylation-specific multiplex ligation-dependent probe amplification (MS-MLPA)-based approach. Mol Med 17(7-8), 726-735.

Pearlman, R., Frankel, W.L., Swanson, B., Zhao, W., Yilmaz, A., Miller, K., Bacher, J., Bigley, C., Nelsen, L., Goodfellow, P.J., Goldberg, R.M., Paskett, E., Shields, P.G., Freudenheim, J.L., Stanich, P.P., Lattimer, I., Arnold, M., Liyanarachchi, S., Kalady, M., Heald, B., Greenwood, C., Paquette, I., Prues, M., Draper, D.J., Lindeman, C., Kuebler, J.P., Reynolds, K., Brell, J.M., Shaper, A.A., Mahesh, S., Buie, N., Weeman, K., Shine, K., Haut, M., Edwards, J., Bastola, S., Wickham, K., Khanduja, K.S., Zacks, R., Pritchard, C.C., Shirts, B.H., Jacobson, A., Allen, B., de la Chapelle, A., Hampel, H., Ohio Colorectal Cancer Prevention Initiative Study, G., 2017. Prevalence and Spectrum of Germline Cancer Susceptibility Gene Mutations Among Patients With Early-Onset Colorectal Cancer. JAMA Oncol 3(4), 464-471.

Pearlman, R., Haraldsdottir, S., de la Chapelle, A., Jonasson, J.G., Liyanarachchi, S., Frankel, W.L., Rafnar, T., Stefansson, K., Pritchard, C.C., Hampel, H., 2019. Clinical characteristics of patients with colorectal cancer with double somatic mismatch repair mutations compared with Lynch syndrome. J Med Genet.

Peltomäki, P., 2001. DNA mismatch repair and cancer. Mutat Res 488(1), 77-85.

Peltomäki, P., 2003. Role of DNA mismatch repair defects in the pathogenesis of human cancer. J Clin Oncol 21(6), 1174-1179.

Peltomäki, P., 2005. Lynch syndrome genes. Fam Cancer 4(3), 227-232.

Peltomäki, P., 2012. Mutations and epimutations in the origin of cancer. Exp Cell Res 318(4), 299-310.

Peltomäki, P., 2014. Epigenetic mechanisms in the pathogenesis of Lynch syndrome. Clin Genet 85(5), 403-412.

Peltomäki, P., 2017. Encyclopedia of Cancer 3rd Edition, 3rd ed.

Peltomäki, P., Aaltonen, L.A., Sistonen, P., Pylkkanen, L., Mecklin, J.P., Jarvinen, H., Green, J.S., Jass, J.R., Weber, J.L., Leach, F.S., et al., 1993. Genetic mapping of a locus predisposing to human colorectal cancer. Science 260(5109), 810-812.

Peng, Y., Croce, C.M., 2016. The role of MicroRNAs in human cancer. Signal Transduct Target Ther 1, 15004.

Pino, M.S., Chung, D.C., 2010. The chromosomal instability pathway in colon cancer. Gastroenterology 138(6), 2059-2072.

Qi, Y., Wang, D., Wang, D., Jin, T., Yang, L., Wu, H., Li, Y., Zhao, J., Du, F., Song, M., Wang, R., 2016. HEDD: the human epigenetic drug database. Database (Oxford) 2016.

Rahman, N., 2014. Realizing the promise of cancer predisposition genes. Nature 505(7483), 302-308.

Renaux-Petel, M., Charbonnier, F., Thery, J.C., Fermey, P., Lienard, G., Bou, J., Coutant, S., Vezain, M., Kasper, E., Fourneaux, S., Manase, S., Blanluet, M., Leheup, B., Mansuy, L., Champigneulle, J., Chappe, C., Longy, M., Sevenet, N., Paillerets, B.B., Guerrini-Rousseau, L., Brugieres, L., Caron, O., Sabourin, J.C., Tournier, I., Baert-Desurmont, S., Frebourg, T., Bougeard, G., 2018. Contribution of de novo and mosaic TP53 mutations to Li-Fraumeni syndrome. J Med Genet 55(3), 173-180.

Riaz, N., Blecua, P., Lim, R.S., Shen, R., Higginson, D.S., Weinhold, N., Norton, L., Weigelt, B., Powell, S.N., Reis-Filho, J.S., 2017. Pan-cancer analysis of bi-allelic alterations in homologous recombination DNA repair genes. Nat Commun 8(1), 857.

Roberts, M.E., Jackson, S.A., Susswein, L.R., Zeinomar, N., Ma, X., Marshall, M.L., Stettner, A.R., Milewski, B., Xu, Z., Solomon, B.D., Terry, M.B., Hruska, K.S., Klein, R.T., Chung, W.K., 2018. MSH6 and PMS2 germ-line pathogenic variants implicated in Lynch syndrome are associated with breast cancer. Genet Med.

Rodriguez-Bigas, M.A., Boland, C.R., Hamilton, S.R., Henson, D.E., Jass, J.R., Khan, P.M., Lynch, H., Perucho, M., Smyrk, T., Sobin, L., Srivastava, S., 1997. A National Cancer Institute Workshop on Hereditary Nonpolyposis Colorectal Cancer Syndrome: meeting highlights and Bethesda guidelines. J Natl Cancer Inst 89(23), 1758-1762.

Rodriguez-Soler, M., Perez-Carbonell, L., Guarinos, C., Zapater, P., Castillejo, A., Barbera, V.M., Juarez, M., Bessa, X., Xicola, R.M., Clofent, J., Bujanda, L., Balaguer, F., Rene, J.M., de-Castro, L., Marin-Gabriel, J.C., Lanas, A., Cubiella, J., Nicolas-Perez, D., Brea-Fernandez, A., Castellvi-Bel, S., Alenda, C., Ruiz-Ponte, C., Carracedo, A., Castells, A., Andreu, M., Llor, X., Soto, J.L., Paya, A., Jover, R., 2013. Risk of cancer in cases of suspected lynch syndrome without germline mutation. Gastroenterology 144(5), 926-932 e921; quiz e913-924.

Romero-Laorden, N., Castro, E., 2017. Inherited mutations in DNA repair genes and cancer risk. Curr Probl Cancer 41(4), 251-264.

Rosenthal, R., McGranahan, N., Herrero, J., Taylor, B.S., Swanton, C., 2016. DeconstructSigs: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. Genome Biol 17, 31.

Rossi, L., Le Frere-Belda, M.A., Laurent-Puig, P., Buecher, B., De Pauw, A., Stoppa-Lyonnet, D., Canlorbe, G., Caron, O., Borghese, B., Colas, C., Delhomelle, H., Chabbert-Buffet, N., Grandjouan, S., Lecuru, F., Bats, A.S., 2017. Clinicopathologic Characteristics of Endometrial Cancer in Lynch Syndrome: A French Multicenter Study. Int J Gynecol Cancer 27(5), 953-960.

Ryan, N.A.J., Evans, D.G., Green, K., Crosbie, E.J., 2017. Pathological features and clinical behavior of Lynch syndrome-associated ovarian cancer. Gynecol Oncol 144(3), 491-495.

Saita, C., Yamaguchi, T., Horiguchi, S.I., Yamada, R., Takao, M., Iijima, T., Wakaume, R., Aruga, T., Tabata, T., Koizumi, K., 2018. Tumor development in Japanese patients with Lynch syndrome. PLoS One 13(4), e0195572.

Samadder, N.J., Smith, K.R., Wong, J., Thomas, A., Hanson, H., Boucher, K., Kopituch, C., Cannon-Albright, L.A., Burt, R.W., Curtin, K., 2017. Cancer Risk in Families Fulfilling the Amsterdam Criteria for Lynch Syndrome. JAMA Oncol 3(12), 1697-1701.

Schmidt, A.Y., Hansen, T.V.O., Ahlborn, L.B., Jonson, L., Yde, C.W., Nielsen, F.C., 2017. Next-Generation Sequencing-Based Detection of Germline Copy Number Variations in BRCA1/BRCA2: Validation of a One-Step Diagnostic Workflow. J Mol Diagn 19(6), 809-816.

Schultheis, A.M., Ng, C.K., De Filippo, M.R., Piscuoglio, S., Macedo, G.S., Gatius, S., Perez Mies, B., Soslow, R.A., Lim, R.S., Viale, A., Huberman, K.H., Palacios, J.C., Reis-Filho, J.S., Matias-Guiu, X., Weigelt, B., 2016. Massively Parallel Sequencing-Based Clonality Analysis of Synchronous Endometrioid Endometrial and Ovarian Carcinomas. J Natl Cancer Inst 108(6), djv427.

Schwarz, J.M., Cooper, D.N., Schuelke, M., Seelow, D., 2014. MutationTaster2: mutation prediction for the deep-sequencing age. Nat Methods 11(4), 361-362.

Sharma, S., Kelly, T.K., Jones, P.A., 2010. Epigenetics in cancer. Carcinogenesis 31(1), 27-36.

Shia, J., 2008. Immunohistochemistry versus microsatellite instability testing for screening colorectal cancer patients at risk for hereditary nonpolyposis colorectal cancer syndrome. Part I. The utility of immunohistochemistry. J Mol Diagn 10(4), 293-300.

Shihab, H.A., Gough, J., Cooper, D.N., Day, I.N., Gaunt, T.R., 2013a. Predicting the functional consequences of cancer-associated amino acid substitutions. Bioinformatics 29(12), 1504-1510.

Shihab, H.A., Gough, J., Cooper, D.N., Stenson, P.D., Barker, G.L., Edwards, K.J., Day, I.N., Gaunt, T.R., 2013b. Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. Hum Mutat 34(1), 57-65.

Shihab, H.A., Gough, J., Mort, M., Cooper, D.N., Day, I.N., Gaunt, T.R., 2014. Ranking non-synonymous single nucleotide polymorphisms based on disease concepts. Hum Genomics 8, 11.

Sorscher, S., 2019. Rationale for evaluating breast cancers of Lynch syndrome patients for mismatch repair gene expression. Breast Cancer Res Treat 178(2), 469-471.

Stoffel, E.M., Koeppe, E., Everett, J., Ulintz, P., Kiel, M., Osborne, J., Williams, L., Hanson, K., Gruber, S.B., Rozek, L.S., 2018. Germline Genetic Features of Young Individuals With Colorectal Cancer. Gastroenterology 154(4), 897-905 e891.

Stratton, M.R., Campbell, P.J., Futreal, P.A., 2009. The cancer genome. Nature 458(7239), 719-724.

Sulonen, A.M., Ellonen, P., Almusa, H., Lepisto, M., Eldfors, S., Hannula, S., Miettinen, T., Tyynismaa, H., Salo, P., Heckman, C., Joensuu, H., Raivio, T., Suomalainen, A., Saarela, J., 2011. Comparison of solution-based exome capture methods for next generation sequencing. Genome Biol 12(9), R94.

Sun, W., Yang, J., 2010. Functional mechanisms for human tumor suppressors. J Cancer 1, 136-140.

Terdiman, J.P., 2005. It is time to get serious about diagnosing Lynch syndrome (hereditary nonpolyposis colorectal cancer with defective DNA mismatch repair) in the general population. Gastroenterology 129(2), 741-744.

Thiel, A., Heinonen, M., Kantonen, J., Gylling, A., Lahtinen, L., Korhonen, M., Kytola, S., Mecklin, J.P., Orpana, A., Peltomaki, P., Ristimaki, A., 2013. BRAF mutation in sporadic colorectal cancer and Lynch syndrome. Virchows Arch 463(5), 613-621.

Thompson, B.A., Spurdle, A.B., Plazzer, J.P., Greenblatt, M.S., Akagi, K., Al-Mulla, F., Bapat, B., Bernstein, I., Capella, G., den Dunnen, J.T., du Sart, D., Fabre, A., Farrell, M.P., Farrington, S.M., Frayling, I.M., Frebourg, T., Goldgar, D.E., Heinen, C.D., Holinski-Feder, E., Kohonen-Corish, M., Robinson, K.L., Leung, S.Y., Martins, A., Moller, P., Morak, M., Nystrom, M., Peltomaki, P., Pineda, M., Qi, M., Ramesar, R., Rasmussen, L.J., Royer-Pokora, B., Scott, R.J., Sijmons, R., Tavtigian, S.V., Tops, C.M., Weber, T., Wijnen, J., Woods, M.O., Macrae, F., Genuardi, M., InSiGht, 2014. Application of a 5-tiered scheme for standardized classification of 2,360 unique mismatch repair gene variants in the InSiGHT locus-specific database. Nat Genet 46(2), 107-115.

Toss, A., Tomasello, C., Razzaboni, E., Contu, G., Grandi, G., Cagnacci, A., Schilder, R.J., Cortesi, L., 2015. Hereditary ovarian cancer: not only BRCA 1 and 2 genes. Biomed Res Int 2015, 341723.

Tung, N., Lin, N.U., Kidd, J., Allen, B.A., Singh, N., Wenstrup, R.J., Hartman, A.R., Winer, E.P., Garber, J.E., 2016. Frequency of Germline Mutations in 25 Cancer Susceptibility Genes in a Sequential Series of Patients With Breast Cancer. J Clin Oncol 34(13), 1460-1468.

Umar, A., Boland, C.R., Terdiman, J.P., Syngal, S., de la Chapelle, A., Ruschoff, J., Fishel, R., Lindor, N.M., Burgart, L.J., Hamelin, R., Hamilton, S.R., Hiatt, R.A., Jass, J., Lindblom, A., Lynch, H.T., Peltomaki, P., Ramsey, S.D., Rodriguez-Bigas, M.A., Vasen, H.F., Hawk, E.T., Barrett, J.C., Freedman, A.N., Srivastava, S., 2004. Revised Bethesda Guidelines for hereditary

nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. J Natl Cancer Inst 96(4), 261-268.

Valle, L., de Voer, R.M., Goldberg, Y., Sjursen, W., Forsti, A., Ruiz-Ponte, C., Caldes, T., Garre, P., Olsen, M.F., Nordling, M., Castellvi-Bel, S., Hemminki, K., 2019. Update on genetic predisposition to colorectal cancer and polyposis. Mol Aspects Med.

Valle, L., Hernandez-Illan, E., Bellido, F., Aiza, G., Castillejo, A., Castillejo, M.I., Navarro, M., Segui, N., Vargas, G., Guarinos, C., Juarez, M., Sanjuan, X., Iglesias, S., Alenda, C., Egoavil, C., Segura, A., Juan, M.J., Rodriguez-Soler, M., Brunet, J., Gonzalez, S., Jover, R., Lazaro, C., Capella, G., Pineda, M., Soto, J.L., Blanco, I., 2014. New insights into POLE and POLD1 germline mutations in familial colorectal cancer and polyposis. Hum Mol Genet 23(13), 3506-3512.

Walsh, T., Casadei, S., Lee, M.K., Pennil, C.C., Nord, A.S., Thornton, A.M., Roeb, W., Agnew, K.J., Stray, S.M., Wickramanayake, A., Norquist, B., Pennington, K.P., Garcia, R.L., King, M.C., Swisher, E.M., 2011. Mutations in 12 genes for inherited ovarian, fallopian tube, and peritoneal carcinoma identified by massively parallel sequencing. Proc Natl Acad Sci U S A 108(44), 18032-18037.

Vasen, H.F., Mecklin, J.P., Khan, P.M., Lynch, H.T., 1991. The International Collaborative Group on Hereditary Non-Polyposis Colorectal Cancer (ICG-HNPCC). Dis Colon Rectum 34(5), 424-425.

Vasen, H.F., Morreau, H., Nortier, J.W., 2001. Is breast cancer part of the tumor spectrum of hereditary nonpolyposis colorectal cancer? Am J Hum Genet 68(6), 1533-1535.

Vasen, H.F., Tomlinson, I., Castells, A., 2015. Clinical management of hereditary colorectal cancer syndromes. Nat Rev Gastroenterol Hepatol 12(2), 88-97.

Vasen, H.F., Watson, P., Mecklin, J.P., Lynch, H.T., 1999. New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative group on HNPCC. Gastroenterology 116(6), 1453-1456.

Watson, P., Riley, B., 2005. The tumor spectrum in the Lynch syndrome. Fam Cancer 4(3), 245-248.

Watson, P., Vasen, H.F.A., Mecklin, J.P., Bernstein, I., Aarnio, M., Jarvinen, H.J., Myrhoj, T., Sunde, L., Wijnen, J.T., Lynch, H.T., 2008. The risk of extra-colonic, extra-endometrial cancer in the Lynch syndrome. Int J Cancer 123(2), 444-449.

Weisenberger, D.J., Siegmund, K.D., Campan, M., Young, J., Long, T.I., Faasse, M.A., Kang, G.H., Widschwendter, M., Weener, D., Buchanan, D., Koh, H., Simms, L., Barker, M., Leggett, B., Levine, J., Kim, M., French, A.J., Thibodeau, S.N., Jass, J., Haile, R., Laird, P.W., 2006. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. Nat Genet 38(7), 787-793.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., Gocayne, J.D., Amanatides, P., Ballew, R.M., Huson, D.H., Wortman, J.R., Zhang, Q., Kodira, C.D., Zheng, X.H., Chen, L., Skupski, M., Subramanian, G., Thomas, P.D., Zhang, J., Gabor Miklos, G.L., Nelson, C., Broder, S., Clark,

A.G., Nadeau, J., McKusick, V.A., Zinder, N., Levine, A.J., Roberts, R.J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Di Francesco, V., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A.E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T.J., Higgins, M.E., Ji, R.R., Ke, Z., Ketchum, K.A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G.V., Milshina, N., Moore, H.M., Naik, A.K., Narayan, V.A., Neelam, B., Nusskern, D., Rusch, D.B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., Yan, C., Yao, A., Ye, J., Zhan, M., Zhang, W., Zhang, H., Zhao, Q., Zheng, L., Zhong, F., Zhong, W., Zhu, S., Zhao, S., Gilbert, D., Baumhueter, S., Spier, G., Carter, C., Cravchik, A., Woodage, T., Ali, F., An, H., Awe, A., Baldwin, D., Baden, H., Barnstead, M., Barrow, I., Beeson, K., Busam, D., Carver, A., Center, A., Cheng, M.L., Curry, L., Danaher, S., Davenport, L., Desilets, R., Dietz, S., Dodson, K., Doup, L., Ferriera, S., Garg, N., Gluecksmann, A., Hart, B., Haynes, J., Haynes, C., Heiner, C., Hladun, S., Hostin, D., Houck, J., Howland, T., Ibegwam, C., Johnson, J., Kalush, F., Kline, L., Koduru, S., Love, A., Mann, F., May, D., McCawley, S., McIntosh, T., McMullen, I., Moy, M., Moy, L., Murphy, B., Nelson, K., Pfannkoch, C., Pratts, E., Puri, V., Qureshi, H., Reardon, M., Rodriguez, R., Rogers, Y.H., Romblad, D., Ruhfel, B., Scott, R., Sitter, C., Smallwood, M., Stewart, E., Strong, R., Suh, E., Thomas, R., Tint, N.N., Tse, S., Vech, C., Wang, G., Wetter, J., Williams, S., Williams, M., Windsor, S., Winn-Deen, E., Wolfe, K., Zaveri, J., Zaveri, K., Abril, J.F., Guigo, R., Campbell, M.J., Sjolander, K.V., Karlak, B., Kejariwal, A., Mi, H., Lazareva, B., Hatton, T., Narechania, A., Diemer, K., Muruganujan, A., Guo, N., Sato, S., Bafna, V., Istrail, S., Lippert, R., Schwartz, R., Walenz, B., Yooseph, S., Allen, D., Basu, A., Baxendale, J., Blick, L., Caminha, M., Carnes-Stine, J., Caulk, P., Chiang, Y.H., Coyne, M., Dahlke, C., Mays, A., Dombroski, M., Donnelly, M., Ely, D., Esparham, S., Fosler, C., Gire, H., Glanowski, S., Glasser, K., Glodek, A., Gorokhov, M., Graham, K., Gropman, B., Harris, M., Heil, J., Henderson, S., Hoover, J., Jennings, D., Jordan, C., Jordan, J., Kasha, J., Kagan, L., Kraft, C., Levitsky, A., Lewis, M., Liu, X., Lopez, J., Ma, D., Majoros, W., McDaniel, J., Murphy, S., Newman, M., Nguyen, T., Nguyen, N., Nodell, M., Pan, S., Peck, J., Peterson, M., Rowe, W., Sanders, R., Scott, J., Simpson, M., Smith, T., Sprague, A., Stockwell, T., Turner, R., Venter, E., Wang, M., Wen, M., Wu, D., Wu, M., Xia, A., Zandieh, A., Zhu, X., 2001. The sequence of the human genome. Science 291(5507), 1304-1351.

Williams, M.J., Werner, B., Barnes, C.P., Graham, T.A., Sottoriva, A., 2016. Identification of neutral tumor evolution across cancer types. Nat Genet 48(3), 238-244.

Wimmer, K., Kratz, C.P., Vasen, H.F., Caron, O., Colas, C., Entz-Werle, N., Gerdes, A.M., Goldberg, Y., Ilencikova, D., Muleris, M., Duval, A., Lavoine, N., Ruiz-Ponte, C., Slavc, I., Burkhardt, B., Brugieres, L., CMMRD, E.U.-C.C.f., 2014. Diagnostic criteria for constitutional mismatch repair deficiency syndrome: suggestions of the European consortium 'care for CMMRD' (C4CMMRD). J Med Genet 51(6), 355-365.

Win, A.K., Jenkins, M.A., Dowty, J.G., Antoniou, A.C., Lee, A., Giles, G.G., Buchanan, D.D., Clendenning, M., Rosty, C., Ahnen, D.J., Thibodeau, S.N., Casey, G., Gallinger, S., Le Marchand, L., Haile, R.W., Potter, J.D., Zheng, Y., Lindor, N.M., Newcomb, P.A., Hopper, J.L., MacInnis, R.J., 2017. Prevalence and Penetrance of Major Genes and Polygenes for Colorectal Cancer. Cancer Epidemiol Biomarkers Prev 26(3), 404-412.

Win, A.K., Lindor, N.M., Jenkins, M.A., 2013. Risk of breast cancer in Lynch syndrome: a systematic review. Breast Cancer Res 15(2), R27.

Win, A.K., Young, J.P., Lindor, N.M., Tucker, K.M., Ahnen, D.J., Young, G.P., Buchanan, D.D., Clendenning, M., Giles, G.G., Winship, I., Macrae, F.A., Goldblatt, J., Southey, M.C., Arnold, J., Thibodeau, S.N., Gunawardena, S.R., Bapat, B., Baron, J.A., Casey, G., Gallinger, S., Le Marchand, L., Newcomb, P.A., Haile, R.W., Hopper, J.L., Jenkins, M.A., 2012. Colorectal and other cancer risks for carriers and noncarriers from families with a DNA mismatch repair gene mutation: a prospective cohort study. J Clin Oncol 30(9), 958-964.

Vogelstein, B., Papadopoulos, N., Velculescu, V.E., Zhou, S., Diaz, L.A., Jr., Kinzler, K.W., 2013. Cancer genome landscapes. Science 339(6127), 1546-1558.

Wooster, R., Bignell, G., Lancaster, J., Swift, S., Seal, S., Mangion, J., Collins, N., Gregory, S., Gumbs, C., Micklem, G., 1995. Identification of the breast cancer susceptibility gene BRCA2. Nature 378(6559), 789-792.

Voutsadakis, I.A., 2019. High Tumor Mutation Burden and Other Immunotherapy Response Predictors in Breast Cancers: Associations and Therapeutic Opportunities. Target Oncol.

Xicola, R.M., Clark, J.R., Carroll, T., Alvikas, J., Marwaha, P., Regan, M.R., Lopez-Giraldez, F., Choi, J., Emmadi, R., Alagiozian-Angelova, V., Kupfer, S.S., Ellis, N.A., Llor, X., 2019. Implication of DNA repair genes in Lynch-like syndrome. Fam Cancer 18(3), 331-342.

Yates, A., Akanni, W., Amode, M.R., Barrell, D., Billis, K., Carvalho-Silva, D., Cummins, C., Clapham, P., Fitzgerald, S., Gil, L., Giron, C.G., Gordon, L., Hourlier, T., Hunt, S.E., Janacek, S.H., Johnson, N., Juettemann, T., Keenan, S., Lavidas, I., Martin, F.J., Maurel, T., McLaren, W., Murphy, D.N., Nag, R., Nuhn, M., Parker, A., Patricio, M., Pignatelli, M., Rahtz, M., Riat, H.S., Sheppard, D., Taylor, K., Thormann, A., Vullo, A., Wilder, S.P., Zadissa, A., Birney, E., Harrow, J., Muffato, M., Perry, E., Ruffier, M., Spudich, G., Trevanion, S.J., Cunningham, F., Aken, B.L., Zerbino, D.R., Flicek, P., 2016. Ensembl 2016. Nucleic Acids Res 44(D1), D710-716.

Yurgelun, M.B., Hampel, H., 2018. Recent Advances in Lynch Syndrome: Diagnosis, Treatment, and Cancer Prevention. Am Soc Clin Oncol Educ Book 38, 101-109.

Yurgelun, M.B., Masciari, S., Joshi, V.A., Mercado, R.C., Lindor, N.M., Gallinger, S., Hopper, J.L., Jenkins, M.A., Buchanan, D.D., Newcomb, P.A., Potter, J.D., Haile, R.W., Kucherlapati, R., Syngal, S., Colon Cancer Family, R., 2015. Germline TP53 Mutations in Patients With Early-Onset Colorectal Cancer in the Colon Cancer Family Registry. JAMA Oncol 1(2), 214-221.

Zehir, A., Benayed, R., Shah, R.H., Syed, A., Middha, S., Kim, H.R., Srinivasan, P., Gao, J., Chakravarty, D., Devlin, S.M., Hellmann, M.D., Barron, D.A., Schram, A.M., Hameed, M., Dogan, S., Ross, D.S., Hechtman, J.F., DeLair, D.F., Yao, J., Mandelker, D.L., Cheng, D.T., Chandramohan, R., Mohanty, A.S., Ptashkin, R.N., Jayakumaran, G., Prasad, M., Syed, M.H., Rema, A.B., Liu, Z.Y., Nafa, K., Borsu, L., Sadowska, J., Casanova, J., Bacares, R., Kiecka, I.J., Razumova, A., Son, J.B., Stewart, L., Baldi, T., Mullaney, K.A., Al-Ahmadie, H., Vakiani, E., Abeshouse, A.A., Penson, A.V., Jonsson, P., Camacho, N., Chang, M.T., Won, H.H., Gross, B.E., Kundra, R., Heins, Z.J., Chen, H.W., Phillips, S., Zhang, H., Wang, J., Ochoa, A., Wills, J., Eubank, M., Thomas, S.B., Gardos, S.M., Reales, D.N., Galle, J., Durany, R., Cambria, R., Abida, W., Cercek, A., Feldman, D.R., Gounder, M.M., Hakimi, A.A., Harding, J.J., Iyer, G., Janjigian, Y.Y., Jordan, E.J., Kelly, C.M., Lowery, M.A., Morris, L.G.T., Omuro, A.M., Raj,

N., Razavi, P., Shoushtari, A.N., Shukla, N., Soumerai, T.E., Varghese, A.M., Yaeger, R., Coleman, J., Bochner, B., Riely, G.J., Saltz, L.B., Scher, H.I., Sabbatini, P.J., Robson, M.E., Klimstra, D.S., Taylor, B.S., Baselga, J., Schultz, N., Hyman, D.M., Arcila, M.E., Solit, D.B., Ladanyi, M., Berger, M.F., 2017. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. Nat Med 23(6), 703-713.

ORIGINAL PUBLICATIONS I-IV